

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 23, 2003, 09:16:28 : Search time 624 Seconds
(without alignments)
8033.424 Million cell updates/sec

Title: US-10-000-151b-1

Perfect score: 1857

Sequence: 1 atgagcagcagctagaggtta.....aagtaatatgncncaaa 1857

Scoring table: IDENTITY_NOC
Gapop 10.0 , Gapext 1.0

Searched: 2552756 segs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N.Geneseq_19Jun03.*

```
1: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:*
2: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:*
3: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:*
4: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:*
5: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:*
6: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:*
7: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:*
8: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:*
9: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:*
10: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:*
11: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:*
12: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:*
13: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:*
14: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:*
15: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:*
16: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:*
17: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:*
18: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:*
19: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:*
20: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:*
21: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:*
22: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:*
23: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:*
24: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:*
25: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA2003.DAT:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1810.6	97.5	1857	24	ABK86572
2	1645.4	88.6	2280	24	AAAD44094
3	1359.8	73.2	1461	24	ABT05463
4	1311.8	70.6	1489	24	ABT05462
5	1217.2	65.5	2362	22	AAH18539
6	899.6	48.4	1165	21	AAA87668
7	894.6	48.2	1160	21	AAA87699
8	725.2	39.1	807	22	AAH06644

9	509.4	27.4	843	24	ABE211754	Human polynucleoti
10	502	27.0	539	21	AAZ58638	Pancreatic cancer
11	350.4	18.9	387	21	AAK89423	Human digestive sy
12	343.2	18.5	469	22	AAFP93488	CDNA encoding SPT
13	336.4	18.1	442	22	AAFP93569	Lung carcinoma CDN
14	334.2	18.0	348	22	AAH98871	Rat EST-derived co
15	258.4	13.9	283	20	AAV87898	EST clone FE311.
16	174.6	9.4	856	22	AAH06371	Human CDNA clone (
17	174.6	9.4	2245	22	AAH16194	Human CDNA sequenc
18	151.8	8.2	546	22	AAH11428	Human CDNA clone (
19	86.8	4.7	10233	23	ABL13164	Drosophila melanog
20	85.2	4.6	3672	23	ABL13165	Drosophila melanog
21	72	3.9	6644	20	AAK33181	Base sequence of t
22	72	3.9	7372	20	AAK33182	Base sequence of t
23	72	3.9	7797	20	AAK33180	Cowpox virus bsr f
24	72	3.9	7996	20	AAK33184	Base sequence of t
25	71.2	3.8	4545	23	ABL10718	Drosophila melanog
26	71	3.8	6668	24	ABL33697	Human immune syste
27	68	3.7	14006	24	ABL33958	Human immune syste
28	67.2	3.6	6292	22	AAK46735	Tumour suppressor
29	65.8	3.5	16033	24	ABL34404	Human immune syste
30	65.2	3.5	8079	24	ABL92313	Chemically treated
31	65	3.5	3683	25	ABE10199	Haematopoietic cel
32	65	3.5	7442	22	AAK46686	Tumour suppressor
33	64.8	3.5	9155	24	ABL32462	Human immune syste
34	64.2	3.5	6898	24	ABH80222	Human chemically m
35	64.2	3.5	9539	22	AAK45347	Chemically pretrea
36	64.2	3.5	9539	24	ABK28180	DNA transcription
37	63.6	3.4	6767	22	AAK46608	Tumour suppressor
38	63.4	3.4	34769	22	AAK46774	Tumour suppressor
39	63	3.4	875	22	AAK46781	Human neuroblastom
40	62.6	3.4	4654	22	AAK46781	Tumour suppressor
41	62.6	3.4	4654	22	AAK46781	Human immune syste
42	62.2	3.3	4554	24	ABL34223	Human immune syste
43	62	3.3	7571	24	ABL32527	Human immune syste
44	61.6	3.3	17934	24	ABL33719	Human immune syste
45	61.6	3.3	819	22	AAK49474	Human neuroblastom
					AAK195036	Human neuroblastom

ALIGNMENTS

RESULT 1	ABK86572	standard: CDNA: 1857 BP.
ID	ABK86572	
AC	ABK86572;	
XX		
DT	24-SEP-2002 (first entry)	
XX		
DE	CDNA encoding human ether-a-go-go related interacting protein KCRLb.	
XX		
KW	Human; ss; gene; human ether-a-go-go related gene; HERG; KCRL; SNP;	
KW	Long QT syndrome; LQT; single nucleotide polymorphism; cardiac arrhythmia;	
KW	potassium channel.	
XX		
OS	Homo sapiens.	
XX		
PH	key	Location/Qualifiers
FT	CDS	1..1422
FT		/tag= a
FT		/product= "KCRLb"
FT	variation	replace (1339,A)
FT		/tag= b
FT		/standard_name= "Single nucleotide polymorphism"
XX		
PN	WO200242735-A2.	
XX		
XX	30-MAY-2002.	
PD		
XX		
XX	30-OCT-2001; 2001WO-US45644.	
PF		
XX		
PR	30-OCT-2000; 2000US-244340P.	

XX (UYVA-) UNIV VANDERBILT.
XX
XX
XX Balser JR, George AL, Roden DM;
XX WPI: 2002-527650/56.
XX P-PSDB; AA099166.
XX
XX Identifying a potassium channel activity modulator for drug design,
XX comprises contacting a compound with a potassium channel and rat
XX cerebellar cDNA library (KCR1) polypeptide, and determining activity
XX
XX Claim 11: Page 154-156; 164pp; English.
XX
XX The invention relates to identifying (M1) a compound that modulates
XX biological activity of a potassium channel (PC), by contacting a
XX compound with a structure comprising a PC polypeptide and a polypeptide
XX cloned from a rat cerebellar cDNA library (KCR1), and determining the
XX activity of the PC polypeptide in the presence and absence of the
XX compound, where a difference in the activities indicates modulation of
XX biological activity of PC. Also included are identifying (M2) a candidate
XX compound that modulates the biological activity of a complex comprising a
XX human ether-a-go-go-related gene (HERG) channel polypeptide and a KCR1
XX polypeptide, identifying (M3) a candidate compound as a modulator of KCR1
XX expression, modulating (M4) PC function in a subject, comprising
XX administering to the subject a substance that provides expression of a
XX KCR1-encoding nucleic acid molecule in a cell or tissue, where modulated
XX PC function is desired, screening (M5) for susceptibility to a drug-
XX induced cardiac arrhythmia in a subject, comprising obtaining a
XX biological sample from the subject and detecting a polymorphism of a KCR1
XX gene in the biological sample from the subject, where the presence of the
XX polymorphism indicates the susceptibility of the subject to a
XX drug-induced cardiac arrhythmia, an oligonucleotide pair, where a first
XX oligonucleotide of the pair hybridizes to a first portion of a KCR1 gene
XX which includes a polymorphism of the KCR1 gene, and the second
XX oligonucleotide of the pair hybridizes to a second portion of the KCR1
XX gene that is adjacent to the first portion and a set of antisense
XX oligonucleotide primers, suitable for amplifying a portion of a KCR1 gene
XX which includes a polymorphism of the KCR1 gene. (M1) is useful for
XX identifying a compound that modulates biological activity of PC,
XX especially HERG, for modulating PC function (i.e. modulating HERG
XX activity) in a mammal, by preparing a composition comprising the
XX compound and administering the composition. The compound is useful for
XX treating or preventing long QT syndrome (LQT) and is useful in drug
XX designing. The present sequence encodes the human KCR1b allele.
XX
XX Sequence 1857 BP; 475 A; 343 C; 334 G; 652 T; 53 other;
XX
XX Query Match 97.5%; Score 1810.6; DB 24; Length 1857;
XX Best Local Similarity 100.0%; Pred. No. 0;
XX Matches 1857; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 1 ATGGCGAGCTAGAGGCTTACTGTTTCGCGCGCTGAGCTGACCTTTTATGTC 60
DB 1 ATGGCGAGCTAGAGGCTTACTGTTTCGCGCGCTGAGCTGACCTTTTATGTC 60
QY 61 TGCCTCTCTTCCTCGGCTTACGCCGGGCGCTGCGAGAGCCCTACATGAGAGATCTTC 120
DB 61 TGCCTCTCTTCCTCGGCTTACGCCGGGCGCTGCGAGAGCCCTACATGAGAGATCTTC 120
QY 121 CACCTGCTCAGGCGGAGCGCTACTGTGAGGCGCATTTCTCCCTTCCCATGGGATCC 180
DB 121 CACCTGCTCAGGCGGAGCGCTACTGTGAGGCGCATTTCTCCCTTCCCATGGGATCC 180
QY 181 ATGATTTACATTAACCTGCTTGTACCTGTGAGTGTGAGTGTGAGTGTGAGTGTG 240
DB 181 ATGATTTACATTAACCTGCTTGTACCTGTGAGTGTGAGTGTGAGTGTGAGTGTG 240
QY 241 TGGATCTTTGGATGTGAGATGTTGTCTGCTCATTTGGATGCTGAGATTTGTTAT 300
DB 241 TGGATCTTTGGATGTGAGATGTTGTCTGCTCATTTGGATGCTGAGATTTGTTAT 300
QY 301 CTTCCTTCAGGTGGCAACTCTTATTACTATTGCTTTTCCACAGGATCAACCC 360

DB 301 CTTCCTTCAGGTGGCAACTCTTATTACTATTGCTTTTCCACAGGATCAACCC 360
QY 361 AGAACAAGGCGCTGAGTATGACAGAGCTGTCGACATTAACCTGAGCATTTT 420
DB 361 AGAACAAGGCGCTGAGTATGACAGAGCTGTCGACATTAACCTGAGCATTTT 420
QY 421 CCAACACTTTATTTTAACTTCCCTTATTAACAGAGAGATCTATGTTTAACT 480
DB 421 CCAACACTTTATTTTAACTTCCCTTATTAACAGAGAGATCTATGTTTAACT 480
QY 481 CTTTGTGATATTTGATGTCTTTATGGAATATATAACTTACGCTTCTTGATTT 540
DB 481 CTTTGTGATATTTGATGTCTTTATGGAATATATAACTTACGCTTCTTGATTT 540
QY 541 TGTGGCTTCATGTTTGGCAACAATATCATCTGGGCTGCTCTGCGAGGAATGTC 600
DB 541 TGTGGCTTCATGTTTGGCAACAATATCATCTGGGCTGCTCTGCGAGGAATGTC 600
QY 601 ATTCGACAAAAGTTAACTGAGGCTTGGAAAAGTGAAGTGAAGAGAGACTT 660
DB 601 ATTCGACAAAAGTTAACTGAGGCTTGGAAAAGTGAAGTGAAGAGAGACTT 660
QY 661 CCACCTATTAAGAGACCACTTGGCAAAATTCGAAAATCTTCAGTTCTTTGGCTTAT 720
DB 661 CCACCTATTAAGAGACCACTTGGCAAAATTCGAAAATCTTCAGTTCTTTGGCTTAT 720
QY 721 TCCATGTCCTTTAAACCTTGTAGTATGCTTTGTGAGTATGAGCTTCTTCTG 780
DB 721 TCCATGTCCTTTAAACCTTGTAGTATGCTTTGTGAGTATGAGCTTCTTCTG 780
QY 781 GGATTTCTGTTTGTCTTGTGATGATGATGATGATGATGATGATGATGATGAT 840
DB 781 GGATTTCTGTTTGTCTTGTGATGATGATGATGATGATGATGATGATGATGAT 840
QY 841 AGTATGAGGCTGCTTCACTTTCCTCACTATTTCTTCTTCTTCTTCTTCTTCTT 900
DB 841 AGTATGAGGCTGCTTCACTTTCCTCACTATTTCTTCTTCTTCTTCTTCTTCTT 900
QY 901 TTTTCTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 960
DB 901 TTTTCTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 960
QY 961 AAACATGGAATTCGTTTGTGATGATGATGATGATGATGATGATGATGATGAT 1020
DB 961 AAACATGGAATTCGTTTGTGATGATGATGATGATGATGATGATGATGATGAT 1020
QY 1021 ACTTATGCTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1080
DB 1021 ACTTATGCTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1080
QY 1081 AGAGTTTTCGAAATATGCAATTCGAAATATGCAATTCGAAATATGCAATTCGAAAT 1140
DB 1081 AGAGTTTTCGAAATATGCAATTCGAAATATGCAATTCGAAATATGCAATTCGAAAT 1140
QY 1141 GGTGAGATATAGTGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1200
DB 1141 GGTGAGATATAGTGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1200
QY 1201 ATATGCTGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1260
DB 1201 ATATGCTGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1260
QY 1261 CTTATGCTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1320
DB 1261 CTTATGCTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1320
QY 1321 CTGAGTTGCTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1380
DB 1321 CTGAGTTGCTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1380
QY 1381 CAGTGGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1440
DB 1381 CAGTGGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1440

Db 1381 CAGTGGCCAAATAGTCAGACATTCAAAGCTTTATGTGTAAATATCATGATATTTTGAA 1440
 QY 1441 CTGTAAAAATGACCTTAATATAGACCAATTTCTACAAAACAGCACTGATAGNGGAA 1500
 Db 1441 CTGTAAAAATGACCTTAATATAGACCAATTTCTACAAAACAGCACTGATAGNGGAA 1500
 QY 1501 CATGGAATTTCTTTAGGTGAGTGTGTCTTCAAAATATACATTAATTTTAAATATAT 1560
 Db 1501 CATGGAATTTCTTTAGGTGAGTGTGTCTTCAAAATATACATTAATTTTAAATATAT 1560
 QY 1561 ATTTTAAACATATGTAGAAATTAAGTGGCAAAAGAACTGGGAAAGCTTAAAGCTCTTC 1620
 Db 1561 ATTTTAAACATATGTAGAAATTAAGTGGCAAAAGAACTGGGAAAGCTTAAAGCTCTTC 1620
 QY 1621 AAAGGCTGATATAGGAAATTAANNGTTTNCGATATCATATGCTCCTNNKNT 1680
 Db 1621 AAAGGCTGATATAGGAAATTAANNGTTTNCGATATCATATGCTCCTNNKNT 1680
 QY 1681 GNTGGCCCTTMCNCAAGCTTGGGAATGKTNNMTGNATAGTTNATTAANCTGGGNT 1740
 Db 1681 GNTGGCCCTTMCNCAAGCTTGGGAATGKTNNMTGNATAGTTNATTAANCTGGGNT 1740
 QY 1741 GCTNNMATNACTNNKNCACCCNNNNMACNATNNNTANNATATTTACAAAGCTC 1800
 Db 1741 GCTNNMATNACTNNKNCACCCNNNNMACNATNNNTANNATATTTACAAAGCTC 1800
 QY 1801 AGGTGATTTCTTGAAGTGAAGTGTCTTAAACATTAAGTAAATATGNGCCCAAA 1857
 Db 1801 AGGTGATTTCTTGAAGTGAAGTGTCTTAAACATTAAGTAAATATGNGCCCAAA 1857

RESULT 2
 AAD44094
 ID AAD44094 standard; cdna; 2280 BP.
 AC AAD44094;

DT 13-DEC-2002 (first entry)
 DE Human transmembrane protein (TMP) -3 cDNA.
 XX

KW Human; transmembrane protein; TMP-1; developmental disorder; epilepsy;
 KW prostaticitis; infertility; neurological disorder; Alzheimer's disease;
 KW anaemia; stroke; cardiovascular disorder; hypertension; atherosclerosis;
 KW gastroenterological disorder; anorexia; Crohn's disease; lipid metabolism;
 KW hypercholesterolaemia; hyperlipidaemia; cell proliferative disorder;
 KW psoriasis; autoimmune disorder; acquired immune deficiency syndrome;
 KW AIDS; cancer; gout; Grave's disease; transgenic; transgenic animal;
 KW gene therapy; antileukemia; anticonvulsant; hypotensive; nocturnal;
 KW neuroprotective; cerebroprotective; antiinflammatory; cytoskeletal;
 KW antihypertoid; gene; ss.

OS Homo sapiens.
 FH key Location/Qualifiers
 FT CDS 94..1515
 FT /*tag= a
 FT /product= "Human TMP-3"

XX W0200234783-A2.
 PD 02-MAY-2002.
 PF 26-OCT-2001; 2001MO-US49670.
 PR 27-OCT-2000; 2000US-244017P.
 PR 22-NOV-2000; 2000US-252855P.
 PR 07-DEC-2000; 2000US-251825P.
 PR 12-DEC-2000; 2000US-255085P.
 XX
 PA (INCY-) INCYTE GENOMICS INC.
 XX
 PI Warren BA, Xu Y, Yue H, Batra S, Burford N, Gandhi AR, Walla NK;

PI Arvitu C, Tang YT, Lu DAM, Duggan BM, Baughn MR, Lee EA, Khan FA;
 PI Nguyen DB, Azimzai Y, Yao MG, Lai PG, Thangavelu K, Ramkumar J;
 PI Tran B, Ding L, Au-Young J;
 XX
 DR WPI; 2002-463354/49.
 XX P-PSDB; AAE26417.
 XX
 PT Novel human transmembrane proteins and polynucleotides useful for
 PT diagnosing, treating or preventing infertility, anemia, hypertension,
 PT anorexia, hypercholesterolemia, cancer, gout, Grave's disease
 PS Claim 75; Page 148-149; 163pp; English.

CC The present invention relates to novel human transmembrane proteins (TMP)
 CC and polynucleotides encoding such proteins. Sequences of the invention
 CC are useful for treating diseases or conditions associated with abnormal
 CC expression of TMP such as disorders of reproduction (e.g. infertility,
 CC prostaticitis), developmental (e.g. anaemia, epilepsy), gastrointestinal
 CC (e.g. anorexia, Crohn's disease), neurological (e.g. Alzheimer's disease,
 CC stroke), lipid metabolism (e.g. hypercholesterolaemia, hyperlipidaemia),
 CC cardiovascular (e.g. atherosclerosis, hypertension), cell proliferative
 CC (e.g. cancer, psoriasis) and autoimmune disorders (e.g. acquired immune
 CC deficiency syndrome (AIDS), gout, Grave's disease). They are useful for
 CC creating knockout humanised animals or transgenic animals to model human
 CC disease. Sequences of the invention are also used in gene therapy. The
 CC present sequence is TMP-3 cDNA.

SO Sequence 2280 BP; 578 A; 490 C; 464 G; 748 T; 0 other;

Query Match 88.6%; Score 1645.4; DB 24; Length 2280;

Best Local Similarity 95.8%; Pred. No. 0;

Matches 1678; Conservative 7; Mismatches 66; Indels 1; Gaps 1;

QY 1 ATGGCGCAGCTAGAGGCTTACTGTTCTGCGCGGCTTACAGCTGCTTTTATGATGCC 60
 Db 94 ATGGCGCAGCTAGAGGCTTACTGTTCTGCGCGGCTTACAGCTGCTTTTATGATGCC 153
 QY 61 TGGCTTCCTCTTCCGCTTCAAGCGGCGCTGCGAGAGCCCTACATGAGCAGATCTTC 120
 Db 154 TGGCTTCCTCTTCCGCTTCAAGCGGCGCTGCGAGAGCCCTACATGAGCAGATCTTC 213
 QY 121 CACCTTCCTCAGCGCGAGCGCTACTGTAGAGGCGATTTCTCCTTTCCAGTAGGATGCC 180
 Db 214 CACCTTCCTCAGCGCGAGCGCTACTGTAGAGGCGATTTCTCCTTTCCAGTAGGATGCC 273
 QY 181 ATGATTTACTACATTTACTGCTGCTGACCGGTGTGAGTGTGAGTGTGCAACCTGCATT 240
 Db 274 ATGATTTACTACATTTACTGCTGCTGACCGGTGTGAGTGTGAGTGTGCAACCTGCATT 333
 QY 241 TGGATCTTGGATGCTGTAACATGTTGCTGCTCATTGGAGTGTCAAGATTTGTTAT 300
 Db 334 TGGATCTTGGATGCTGTAACATGTTGCTGCTCATTGGAGTGTCAAGATTTGTTAT 393
 QY 301 CTTCCTCTCAGTGTGGCAACTCTTACTATTTGCTTTCCCAAGATGACAAACC 360
 Db 394 CTTCCTCTCAGTGTGGCAACTCTTACTATTTGCTTTCCCAAGATGACAAACC 453
 QY 361 AGAACAAGGCTGCTCAAGTATCGAGAGTGTGTCACATTAACACATGACATGATT 420
 Db 454 AGAACAAGGCTGCTCAAGTATCGAGAGTGTGTCACATTAACACATGACATGATT 513
 QY 421 CCAACACTTTATTTTAACTCTTATTTATACAGAAAGCAGATCTATGTTTACT 480
 Db 514 CCAACACTTTATTTTAACTCTTATTTATACAGAAAGCAGATCTATGTTTACT 573
 QY 481 CTTTGGCATTTGATGCTGTTTANAGAAATCAATAAACTCAGCTCCCTGATTT 540
 Db 574 CTTTGGCATTTGATGCTGTTTANAGAAATCAATAAACTCAGCTCCCTGATTT 633
 QY 541 TGTGGCTTATGTTTGGCAACAATATCATTTGGGCTGTCTTGTGAGGGAATGTC 600
 Db 634 TGTGGCTTATGTTTGGCAACAATATCATTTGGGCTGTCTTGTGAGGGAATGTC 693

```

QY 601 ATTGCACAAAAGTTAACTGAGCTTGGAAAACTGACTACAAAGAAAGACAGACTT 660
Db 694 ATTGCACAAAAGTTAAAGGAGAGCTTGGAAAACTGAGCTACAAAGAAAGACAGACTT 753
QY 661 CCACCTATTAAAGAGCACTTGGCAATTCAGAAAAATCTTCAGTTCTTTGGCTTAT 720
Db 754 CCACCTATTAAAGAGCACTTGGCAATTCAGAAAAATCTTCAGTTCTTTGGCTTAT 813
QY 721 TCCAGTCCCTTTAAAACTGAGTATGCTTTCTGTTGACTTGGCCCTACATCCTCTG 780
Db 814 TCCAGTCCCTTTAAAACTGAGTATGCTTTCTGTTGACTTGGCCCTACATCCTCTG 873
QY 781 GGATTTCTGTTTGGCTTTGCTTTGTAAGTAAAGTGAATGTTATTTGGCATGAGT 840
Db 874 GGATTTCTGTTTGGCTTTGCTTTGTAAGTAAAGTGAATGTTATTTGGCATGAGT 933
QY 841 AGTCATAGAGCCTGCTTCACTTTTCCACACTATTCATCTTTTTCATTTACTCTCTT 900
Db 934 AGTCATAGAGCCTGCTTCACTTTTCCACACTATTCATCTTTTTCATTTACTCTCTT 993
QY 901 TTTTCTTCCCTCCTCCTCCTCCTCCTAGCAAAATTAAGACTTTTCTTCTTGGTTGG 960
Db 994 TTTTCTTCCCTCCTCCTCCTCCTCCTAGCAAAATTAAGACTTTTCTTCTTGGTTGG 1053
QY 961 AAACATGGAATTCGTTTGGTGGTTACCTTACCTCTCTGTTTGAAGTGAATTC 1020
Db 1054 AAACATGGAATTCGTTTGGTGGTTACCTTACCTCTCTGTTTGAAGTGAATTC 1113
QY 1021 ACTTATGCTCAATTAATCTTGTGACGACACAAATGACATTAATCTTCTATGTGTGAAA 1080
Db 1114 ACTTATGCTCAATTAATCTTGTGACGACACAAATGACATTAATCTTCTATGTGTGAAA 1173
QY 1081 AGAGTTTTCACAAATGACATTCCTGAAATTTGTTAGTTCAGCCCTATATTTGCT 1140
Db 1174 AGAGTTTTCACAAATGACATTCCTGAAATTTGTTAGTTCAGCCCTATATTTGCT 1233
QY 1141 GGTGGAGTATAGCTGACTCATTTGAAAATCAAGCAATTTTGGAAATTAATGTTTTC 1200
Db 1234 GGTGGAGTATAGCTGACTCATTTGAAAATCAAGCAATTTTGGAAATTAATGTTTTC 1293
QY 1201 ATATGCTTGTTCATGTTTATAGTTCCTCAGAAATGCTGGAATTCGTTACTTATTTTA 1260
Db 1294 ATATGCTTGTTCATGTTTATAGTTCCTCAGAAATGCTGGAATTCGTTACTTATTTTA 1353
QY 1261 CCTTATGCTATTTATAGCTTAAACATTAACCTGCTCCACATCAGACTGTTGTTGAAA 1320
Db 1354 CCTTATGCTATTTATAGCTTAAACATTAACCTGCTCCACATCAGACTGTTGTTGAAA 1413
QY 1321 CTGAGTGTGATGCAATTTGTAATTTGCAATCTTTTACATCTTTTGAACAGACTTTT 1380
Db 1414 CTGAGTGTGATGCAATTTGTAATTTGCAATCTTTTGAACAGACTTTT 1473
QY 1381 CAGTGGCCAAATAGTGCAGACATTCAAAGTTTATGTTGTAATTCAGTGTATTTTGGAA 1440
Db 1474 CAGTGGCCAAATAGTGCAGACATTCAAAGTTTATGTTGTAATTCAGTGTATTTTGGAA 1533
QY 1441 CTGTAAAAATGAGCTTATATA-TAGACCACTTTCAGCAAGAAAGCACTGAAATAGAGAAA 1499
Db 1534 CTGTAAAAATGAGCTTATATA-TAGACCACTTTCAGCAAGAAAGCACTGAAATAGAGAAA 1593
QY 1500 ACATGGAATTTCTTTTATGAGTGCAGTGTGCTTCAAAATTAATTAATTTTATATA 1559
Db 1594 ACATGGAATTTCTTTTATGAGTGCAGTGTGCTTCAAAATTAATTAATTTTATATA 1653
QY 1560 TATTTTAAACATATGTAAGAAATTAAGTGCAGAAAGAAAGCTTAAAGACTGCTT 1619
Db 1654 TATTTTAAACATATGTAAGAAATTAAGTGCAGAAAGAAAGCTTAAAGACTGCTT 1713
QY 1620 CAANAAGCTGAATTAATGGAAGAAATTAANWNTTNCAGATATCTCAATGCGCTNNKNA 1679
Db 1714 CAANAAGCTGAATTAATGGAAGAAATTAANWNTTNCAGATATCTCAATGCGCTNNKNA 1773
QY 1680 TGNATGCCCTTMCANAGCTTGGCAATGKTTNNWNTGNATAGTTNATTAANCTGGNN 1739

```

```

Db 1774 TGTGGCCCTCTAAAGCTTGGGATGTTTGTATATACAGATTATTAATACTGGGTA 1833
QY 1740 TGTCTNNMATNA 1751
Db 1834 TGTCTCAAAA 1845

RESULT 3
ABT05463
ID ABT05463 standard; DNA; 1461 BP.
XX
XX ABR05463;
XX
XX 11-OCT-2002 (first entry)
XX
XX DNA of NOVX 9b SEQ ID No 21.
DE
DE Cytostatic; antidiabetic; anorectic; metabolic; nootropic; antilipemic;
KW neuroprotective; antiparkinsonian; anticonvulsant; cerebroprotective;
KW tranquilliser; neuroleptic; antidiabetic; antilucer; antiinflammatory;
KW anti-HIV; antiallergic; antirheumatic; antiarthritic; NOVX; diabetes;
KW metabolic disorder; obesity; infectious disease; Alzheimer's disease;
KW anorexia; neurodegenerative disorder; Parkinson's disease; obesity;
KW immune disorder; haematopoietic disorder; dyslipidaemia; chronic disease;
KW metabolic syndrome X; wasting disorder; cancer; neurological disorder;
KW epilepsy; stroke; mental disorder; schizophrenic disorders; goiter;
KW vesicular transport; cystic fibrosis; gastrointestinal disorder;
KW diabetes mellitus; ulcerative colitis; AIDS; allergic reaction;
KW multiple sclerosis; rheumatoid arthritis; transgenic animal;
KW gene therapy; gene; ds.
XX
XX OS Unidentified.
XX
XX PN W0200246409-A2.
XX
XX PD 13-JUN-2002.
XX
XX PF 06-DEC-2001; 2001WO-0546586.
XX
XX PR 06-DEC-2000; 2000US-251660P.
PR 12-DEC-2000; 2000US-255029P.
PR 08-JAN-2001; 2001US-260326P.
PR 24-JAN-2001; 2001US-263800P.
PR 20-FEB-2001; 2001US-269943P.
PR 24-APR-2001; 2001US-286183P.
PR 20-AUG-2001; 2001US-313627P.
PR 12-SEP-2001; 2001US-318712P.
XX
XX PA (CURA-) CURAGEN CORP.
XX
XX PI Guo X, Li L, Paturajan M, Shinkets RA, Casman SJ, Malyankar UM;
PI Tcherev VT, Vernet CAM, Spytek KA, Shenoy SG, Alsobrook JP;
PI Edinger S, Peyman JA, Stone DJ, Ellerman K, Gangolli EA;
PI Boloid FL, Colman SD, Eisen AJ, Liu X, Padigaru M, Spederna SK;
PI zerhusen BD;
XX
XX WP: 2002-547774/58.
XX
XX DR P-PSDB; ABJ04650.
XX
XX Novel isolated polypeptide, designated NOVX, useful for treating or
PT preventing cancer, diabetes, obesity, dyslipidemia, anorexia, and
PT metabolic, neurodegenerative, immune and hematopoietic disorders -
XX
XX Claim 9; Page 84; 421pp; English.
XX
XX The invention relates to an isolated polypeptide, designated NOVX,
CC comprising a sequence fully defined in the specification. The isolated
CC protein, its encoding polynucleotide or an antibody created from the
CC protein is useful in the manufacture of a medicament for treating a
CC syndrome associated with a human disease, preferably a NOVX-associated
CC disorder, or for treating or preventing a NOVX-associated disorder in a
CC subject, preferably human. The isolated protein, its encoding

```


Db 781 TCCTCAGAAATCGCTGGAATTCCTTACTTATTTTACTTATGTCATTTATAGCTTAA 840
 QY 1284 CATACTCTGCCTCCACATCCAGACTTGTGTGACACTGAGTTCATGCAATTTGTTAA 1343
 Db 841 CATACCTCTGCCTCCACATCCAGACTTGTGTGACACTGAGTTCATGCAATTTGTTAA 900
 QY 1344 TTTCATACCTTTTACATCTTTCTTGACACAGACTTTTACAGGCGCAATATGTCAGACAT 1403
 Db 901 TTTCATACCTTTTACATCTTTCTTGACACAGACTTTTACAGGCGCAATATGTCAGACAT 960
 QY 1404 TCAAGGTTTATGTCATATGATATGATATTTTGAAGTGAATAAGTACTATTA-T 1462
 Db 961 TCAAGGTTTATGTCATATGATATGATATTTTGAAGTGAATAAGTACTATTA-T 1020
 QY 1463 AGACCATTTTACAAAGAACACTGAATAGGNGAAGAACATGATTTCTTTTAGTGCA 1522
 Db 1021 AGACCATTTTACAAAGAACACTGAATAGGNGAAGAACATGATTTCTTTTAGTGCA 1080
 QY 1523 GTGGTGTCTTCAAAATTCATTTAGTTTATATATATATTTTAAACATATGTAAGAAAT 1582
 Db 1081 GTGGTGTCTTCAAAATTCATTTAGTTTATATATATATTTTAAACATATGTAAGAAAT 1140
 QY 1583 TAAGTGGCAAGAACTGGGAAGCTTAGACCTGCTTCAAAAGCCTGAATATGGGAAA 1642
 Db 1141 TAAGTGGCAAGAACTGGGAAGCTTAGACCTGCTTCAAAAGCCTGAATATGGGAAA 1200
 QY 1643 TAAANNGTTTNCAGATATCTCATATCGCTCNCNNKATAGTGGCCCTTMCANAGCTTG 1702
 Db 1201 TAAATTTGTTTTCAGATATCTCATATCTCTCATATAGTTGGCCCTTAAAGACCTTG 1260
 QY 1703 GAATGKTNNMNTGATATAGTTTAAATGAGTGGNNNTGCTNNMNTACT 1753
 Db 1261 GAATGTTTGTATGACAAAGTTTAAACTGGGTATGCTCATATTTACT 1311
 RESULT 6
 AAA87668
 ID AAA87668 standard: cDNA: 1165 BP.
 AC AAA87668;
 XX
 DT 04-DEC-2000 (first entry)
 XX
 DE Human secreted protein gene 3 SEQ ID NO:13.
 XX
 KW Human; secreted protein; immunosuppressive; immunostimulant; nootropic;
 KW antiinflammatory; cardiant; valineray; antilicer; anticonvulsant;
 KW antiparkinsonian; neuroprotective; antiviral; antibacterial; cytostatic;
 KW antiparasitic; thrombolytic; anticoagulant; antiarteriosclerotic;
 KW gene therapy; vaccine; chemotaxis-modulator; angiogenesis-modulator;
 KW cancer; immune system disorder; hyperproliferative disorder; infection;
 KW cardiovascular disorder; neurological disease; wound healing; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200043495-A2.
 XX
 PD 27-JUL-2000.
 XX
 PF 18-JAN-2000; 2000WO-US00903.
 XX
 PR 19-JAN-1999; 99US-0116330.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 PI Rosen CA, Ruben SM, Ebner R, Young PE, Ni J, Moore PA;
 PI Komatsoulis G, Birse CE; Young PE, Ni J, Moore PA;
 XX WPI: 2000-499225/44.
 DR P-PSDB: AAB25667.
 XX
 PT New isolated polynucleotide encoding a secreted protein useful for
 preventing, treating or ameliorating a medical condition -

XX
 PS Claim 1; Page 380; 451pp; English.
 XX
 CC The polynucleotide sequences given in AAB87666 to AAA87708 encodes the
 CC human secreted proteins given in AAB25665 to AAB25755. Human secreted
 CC proteins have activities based on the tissues and cells the genes are
 CC expressed in. Examples of activities include: immunosuppressive;
 CC immunostimulant; antiinflammatory; cardiant; valineray; antilicer;
 CC nootropic; antiviral; anticonvulsant; antiparkinsonian; neuroprotective;
 CC antibacterial; antiparasitic; thrombolytic; anticoagulant; antiarteriosclerotic and cytostatic. The secreted proteins and their
 CC polynucleotides can be used in gene therapy and as vaccines.
 CC chemotaxis-modulators and angiogenesis-modulators. The human secreted
 CC proteins and polynucleotides can be used for diagnosing (the
 CC susceptibility to) a pathological condition by determining the presence
 CC or absence of a mutation in the polynucleotide or determining the
 CC presence or amount of expression of the protein. The polynucleotides and
 CC proteins can also be used in the treatment and diagnosis of cancer,
 CC diseases of the immune system, hyperproliferative disorders,
 CC cardiovascular disorders and neurological disease. They can also be used
 CC to promote wound healing and to fight infection. AAA87657 to AAA87665 and
 CC AAB25664 represent sequences used in the exemplification of the present
 CC invention.
 XX
 SQ Sequence 1165 BP; 290 A; 262 C; 228 G; 385 T; 0 other:
 Query Match 48.4%; Score 899.6; DB 21; Length 1165;
 Best Local Similarity 97.4%; Pred. No. 2.4e-205;
 Matches 914; Conservative 0; Mismatches 24; Indels 0; Gaps 0;
 QY 1 ATGGCGCAGCTAGAGGTTACTGTTCTCGGCGCCCTTGACCTGACTTTTATAGTGC 60
 Db 136 ATGGCGCAGCTAGAGGTTACTGTTCTCGGCGCCCTTGACCTGACTTTTATAGTGC 195
 QY 61 TGCCCTCCTTCCTCCGCTTACGCGGCGGCGGAGAGCCCTACATGAGAGATCTTC 120
 Db 196 TGCCCTCCTTCCTCCGCTTACGCGGCGGCGGAGAGCCCTACATGAGAGATCTTC 255
 QY 121 CACCTGCCTCAGGCGCAGCGCTACTGTGAGGGCCATTTCTCCCTTCCAGTGGAGATCC 180
 Db 256 CACCTGCCTCAGGCGCAGCGCTACTGTGAGGGCCATTTCTCCCTTCCAGTGGAGATCC 315
 QY 181 ATGATTTACTATTAATCTGCTGCTTACTGCTGTGACAGTGGATGTTAACTGCAAT 240
 Db 316 ATGATTTACTATTAATCTGCTGCTTACTGCTGTGACAGTGGATGTTAACTGCAAT 375
 QY 241 TGGATCTTTGATGATGTCGAACATGTTGCTGCATCTGGAGTGCATGATTTGTTAT 300
 Db 376 TGGATCTTTGATGATGTCGAACATGTTGCTGCATCTGGAGTGCATGATTTGTTAT 435
 QY 301 CTCTCTTCACTGTTGGCACTTCTATTTACTATATTTGCTTTCCAAAGGTACAAACC 360
 Db 436 CTCTCTTCACTGTTGGCACTTCTATTTACTATATTTGCTTTCTGCAAGGTACAAACC 495
 QY 361 AGAAACAAGGCTGCTCAAGTATCCAGAGATCTTTGTCACATTTAACTAGCACTATTT 420
 Db 496 AGAAACAAGGCTGCTCAAGTATCCAGAGATCTTTGTCACATTTAACTAGCACTATTT 555
 QY 421 CCAACACTTTTATTTTATTTTAACTTCTTATTTATACAGAGAGATCTATTTTAACT 480
 Db 556 CCAACACTTTTATTTTATTTTAACTTCTTATTTATACAGAGAGATCTATTTTAACT 615
 QY 481 CTTTGTGATATTTGATGATGCTTTATGGAATCATAAACTTACAGCTTCTTGATTT 540
 Db 616 CTTTGTGATATTTGATGATGCTTTATGGAATCATAAACTTACAGCTTCTTGATTT 675
 QY 541 TGTGGCTTCATGTTTGGCAAAATATCATCTGAGGCTGCTTCTGTGACGGAATGTC 600
 Db 676 TGTGGCTTCATGTTTGGCAAAATATCATCTGAGGCTGCTTCTGTGACGGAATGTC 735
 QY 601 ATTGCACAAAAGTTAATGAGGCTTGGAAAAGTACAGCTACAAAGGAAGAGACAGCTT 660
 Db 736 ATTGCACAAAAGTTAATGAGGCTTGGAAAAGTACAGCTACAAAGGAAGAGACAGCTT 795

QY 841 AGTCATGAGCCTGCTTCATTTTCCCTCACTATCTACTTTTTCATTTACTCTCTT 900
DB 968 AGTCATGAGCCTGCTTCATTTTCCCTCACTATCTACTTTTTCATTTACTCTCTT 1027
QY 901 TTTTCCTTTCCATCTCCCTGCTCCCTAGCAAAATTA 938
DB 1028 TTTTCCTTTCCATCTCTCTGCTCAACAATAATAA 1065

RESULT 8
AAH06644
ID AAH06644 standard; cDNA: 807 BP.
AC AAH06644;
XX
XX 26-JUN-2001 (first entry)
DE Human CDNA clone (5'-primer) SEQ ID NO:3479.
XX
XX Human; primer: detection; diagnosis; antisense therapy; gene therapy; ss.
XX Homo sapiens.
XX EP1074617-A2.
XX 07-FEB-2001.
XX 28-JUL-2000; 2000EP-0116126.
XX 29-JUL-1999; 99JP-0248036.
PR 27-AUG-1999; 99JP-0300253.
PR 11-JAN-2000; 2000JP-0118776.
PR 02-MAY-2000; 2000JP-0183767.
PR 09-JUN-2000; 2000JP-0241899.
XX
XX (HELI-) HELIX RES INST.
PA
XX Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX
XX WPI: 2001-318749/34.
XX
XX Primer sets for synthesizing polynucleotides, particularly the 5602
PT full-length cDNAs defined in the specification, and for the detection
PT full-length cDNAs -
XX
XX Claim 1; SEQ ID 3479; 2537pp + CD ROM; English.
PS
XX The present invention describes primer sets for synthesizing 5602
CC full-length cDNAs defined in the specification, where a primer set
CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary
CC to the complementary strand of a polynucleotide which comprises one of
CC the 5602 nucleotide sequences defined in the specification, where the
CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC of an oligonucleotide comprising a sequence complementary to the
CC complementary strand of a polynucleotide which comprises a 5'-end
CC sequence and an oligonucleotide comprising a sequence complementary to a
CC polynucleotide which comprises a 3'-end sequence, where the
CC oligonucleotide comprises at least 15 nucleotides and the combination of
CC the 5'-end sequence/3'-end sequence is selected from those defined in
CC the specification. The primer sets can be used in antisense therapy and
CC in gene therapy. The primers are useful for synthesizing polynucleotides,
CC particularly full-length cDNAs. The primers are also useful for the
CC detection and/or diagnosis of the abnormality of the proteins encoded by
CC the full-length cDNAs. The primers allow obtaining of the full-length
CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to
CC AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632
CC represent oligonucleotides, all of which are used in the exemplification
CC of the present invention.
XX
XX Sequence 807 BP; 200 A; 140 C; 151 G; 312 T; 4 other;

Query Match 39.1%; Score 725.2; DB 22; Length 807;
Best Local Similarity 95.4%; Pred. No. 1.2e-163;
Matches 768; Conservative 0; Mismatches 34; Indels 3; Gaps 2;
QY 444 CCTTATATATACAGACAGACATCATGTTTTTACTCTTTTGTGATATTTGATGTCCT 503
DB 1 CCTTATATATACAGACAGACATCATGTTTTTACTCTTTTGTGATATTTGATGTCCT 60
QY 504 TTATGGAATCATTAATAACTTCAGCCTTCCTTGATTTGTGAGCTTCATGTTGGCAAC 563
DB 61 TTATGGAATCATTAATAACTTCAGCCTTCCTTGATTTGTGAGCTTCATGTTGGCAAC 120
QY 564 AAATATCATCTGGGCTGCTTCTGTCAGGGAATGTCATTCACAAAAGTTAACTGAGC 623
DB 121 AAATATCATCTGGGCTGCTTCTGTCAGGGAATGTCATTCACAAAAGTTAACTGAGC 180
QY 624 TTGGAATACTGAGCTACAAAAGGAAGACAGACTTCCACTATTAAGACCATTTGC 683
DB 181 TTGGAATACTGAGCTACAAAAGGAAGACAGACTTCCACTATTAAGACCATTTGC 240
QY 684 AGAATTCAGAAAATTTCTTCACTTCTTTGAGCTTATTCATGTCCTTAAAACTTGAG 743
DB 241 AGAATTCAGAAAATTTCTTCACTTCTTTGAGCTTATTCATGTCCTTAAAACTTGAG 300
QY 744 TATGCTTTTCTGTTGACTTGACCCTACATCCCTTGGAATTTCTGTTGCTTTTGT 803
DB 301 TATGCTTTTCTGTTGACTTGACCCTACATCCCTTGGAATTTCTGTTGCTTTTGT 360
QY 804 AGTATGTTAATGTTGAATTTGTTATGCGGATGAGATGATGAGAGCCTGCTTCATT 863
DB 361 AGTATGTTAATGTTGAATTTGTTATGCGGATGAGATGATGAGAGCCTGCTTCATT 420
QY 864 TCCCTCACTATTTCACTTTTTCATTTACTCTCTTTTCTTTCTTCTCATCTCCCTGTC 923
DB 421 TCCCTCACTATTTCACTTTTTCATTTACTCTCTTTTCTTTCTTCTCATCTCCCTGTC 480
QY 924 TCCCTCACTATTTCACTTTTTCATTTACTCTCTTTTCTTTCTTCTCATCTCCCTGTC 983
DB 481 TCCCTCACTATTTCACTTTTTCATTTACTCTCTTTTCTTTCTTCTCATCTCCCTGTC 540
QY 984 GGTACCTTACTGCTGCTGCTTTTATGTTGGAATTCATATGCTCAATAACTTCTCT 1043
DB 541 GGTACCTTACTGCTGCTGCTTTTATGTTGGAATTCATATGCTCAATAACTTCTCT 600
QY 1044 AGCAGACATATGACATTTACTTCTATGTTGGAAGAGTTTTCAAAGATATGCAAT 1103
DB 601 AGCAGACATATGACATTTACTTCTATGTTGGAAGAGTTTTCAAAGATATGCAAC 660
QY 1104 TCTGAAATATTTGTTAGTTCAGCCTATATATTTGCTGTTGAGATATAGCTACTATT 1163
DB 661 TGT-AAATATTTGGAATTTCCAGCCTATATATTTGCTGTTGAGATATAGCTACTATT 719
QY 1164 GAATTCAGCAATTTTGTGGAATTTATGTTTTCATATGCTT -GTTCAATGTTATA 1221
DB 720 GAATTCAGCAATTTTGTGGAATTTATGTTTTCATATGCTTCAATGCTATATAG 779
QY 1222 GTTCCTCAGAAACTGCTGAATTTG 1246
DB 780 GTTCCTCAGAAACTGCTGAATTTG 804

RESULT 9
AB211754
ID AB211754 standard; cDNA: 843 BP.
XX
XX AB211754;
AC
XX 20-JAN-2003 (first entry)
DE Human polynucleotide SEQ ID NO 636.
XX
XX Human; genome mapping; gene therapy; food supplement; virus; fungus;

KW cell-proliferative disorder; neurodegenerative disease; bacterial;
 KW Parkinson's disease; Alzheimer's disease; autoimmune disease;
 KW multiple sclerosis; diabetes; genetic disorder; wound; burn; infection;
 KW arthritis; cytostatic; immunomodulator; nootropic; neuroprotective;
 KW antiparkinsonian; antidiabetic; immunosuppressive; dermatological;
 KW hemostatic; vulnery; fungicide; antibacterial; virucide; protozoacide;
 KW antithrilitic; gene; ss.
 XX Homo sapiens.
 OS
 PN NO200270539-A2.
 XX
 PD 12-SEP-2002.
 XX
 PF 05-MAR-2002; 2002MO-US05095.
 XX
 PR 05-MAR-2001; 2001US-0799451.
 XX
 PA (HYSE-) HYSEQ INC.
 PI Tang YT, Zhou P, Goodrich RW, Asundi V, Zhang J, Zhao QA, Ren F;
 PI Xue AJ, Yang Y, Ma Y, Yamazaki V, Chen R, Wang Z, Ghosh M;
 PI Wehrman T, Wang J, Wang D, Dimanac RT;
 DR WPI; 2002-759812/82.
 DR P-PSDB; ABP69537.
 XX
 PT New polynucleotides comprising sequences assembled from expressed
 PT sequence tags (ESTs), useful for treating cell-proliferative,
 PT neurodegenerative, autoimmune, genetic, myeloid or lymphoid, or
 PT platelet or coagulation disorders -
 XX
 PS Claim 1; SEQ ID NO 636; 1012pp + Sequence Listing; English.
 XX
 CC The invention relates to an isolated polynucleotide (I) comprising a
 CC nucleotide sequence selected from any of 948 sequences
 CC (AB211119-AB212066) or their mature protein coding portion, active domain
 CC coding protein or complementary sequences. The polynucleotides are useful
 CC for identifying expressed genes or for physical mapping of human genome.
 CC The encoded polypeptides (ABP68902-ABP69849) are useful as molecular
 CC weight markers, as a food supplement, for generating antibodies, in
 CC medical imaging, screening and diagnostic assays and for treating
 CC cell-proliferative disorders (cancer), neurodegenerative diseases
 CC (Parkinson's or Alzheimer's disease), autoimmune diseases (multiple
 CC sclerosis, diabetes, lupus) genetic disorders, myeloid or lymphoid
 CC disorders, platelet or coagulation disorders, wound, burns, incision,
 CC ulcers, liver or lung fibrosis, infections (bacterial, viral, fungal,
 CC parasitic), arthritis, etc.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 CC
 XX
 SQ Sequence 843 BP; 178 A; 209 C; 191 G; 265 T; 0 other;
 Query Match 27.4%; Score 509.4; DB 24; Length 843;
 Best Local Similarity 95.6%; Pred. No. 5,7e-112;
 Matches 567; Conservative 0; Mismatches 21; Indels 5; Gaps 4;
 QY 1 ATGGCCGAGCTAAGGGTACTGTTCTGCGCGCCCTTGAGCTGACCTTTTATGTC 60
 DB 252 ATGGCCGAGCTAAGGGTACTGTTCTGCGCGCCCTTGAGCTGACCTTTTATGTC 311
 QY 61 TGGCTCTCTTTCGCGCTTACGCGGCGCTGAGAGCCCTACATGAGCGAGTCTTC 120
 DB 312 TGGCTCTCTTTCGCGCTTACGCGGCGCTGAGAGCCCTACATGAGCGAGTCTTC 371
 QY 121 CACCTGCTCAGCGGAGCGCTACTGAGAGGCAATTTCCCTTCCAGTGGATCCC 180
 DB 372 CACCTGCTCAGCGGAGCGCTACTGAGAGGCAATTTCCCTTCCAGTGGATCCC 431
 QY 181 ATGATTACTACTTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 240
 DB 432 ATGATTACTACTTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 491

QY 241 TGGATCTTTGGATGATGCTGAACATGTTGCTGCTCCATTTGGATGCTCAGATTTGTTAAT 300
 DB 492 TGGATCTTTGGATGATGCTGAACATGTTGCTGCTCCATTTGGATGCTCAGATTTGTTAAT 551
 QY 301 CTCTCTTTCAGTGTGGCAACTCTCTATTTACTATTTGCTTTTCCCAAGTATCAACC 360
 DB 552 CTCTCTTTCAGTGTGGCAACTCTCTATTTACTATTTGCTTTTCCCAAGTATCAACC 611
 QY 361 AGAACAAGCTGCTCCTAATATTCAGAGAGTCTTGCAACATTACACTAG-CAGATTT 419
 DB 612 AGAACAAGCTGCTCCTAATATTCAGAGAGTCTTGCAACATTACACTAG-CAGATTT 671
 QY 420 TCCACACCTTTATTTTAACTTCCCTTTATTTATTCAGAGAGAGTCTTATGTTTATAC 479
 DB 672 TCCACACCTTTATTTTAACTTCCCTTTATTTATTCAGAGAGAGTCTTATGTTTATAC 731
 QY 480 TCTTTTGCATATTTGATGCTCTTTATGGAATCATMAAATTCACCTTCTTGATTT 539
 DB 732 TCTTTTGCATATTTGATGCTCTTTATGGAATCATMAAATTCACCTTCTTGATTT 790
 QY 540 TTGTGCTTATGTTTGGCAACAATATATCATCTGGGCTGCTTCTGTGCGAG 592
 DB 791 TTGT-GCTTCATGTTCCGG--AAACAATAATCATCTGGGCTGCTTCTGTGCGG 840

RESULT 10
 AA258638
 ID AA258638 standard; DNA; 539 BP.
 XX
 AC AA258638;
 XX
 DT 11-APR-2000 (first entry)
 XX
 DE Pancreatic cancer and dysplasia up-regulated polynucleotide #5.
 XX
 KW Pancreatic cancer; dysplasia; forensic; genetic analysis; cancer; bone;
 KW brain tumour; breast cancer; endocrine system cancer; gastrointestinal;
 KW genitourinary; gynecological; leukemia; Hodgkin's lymphoma; pediatric;
 KW skin; urinary tract; human; ds.
 XX
 OS Homo sapiens.
 XX
 PN WO9967386-A2.
 XX
 PD 29-DEC-1999.
 XX
 PF 22-JUN-1999; 99MO-US14036.
 XX
 PR 23-JUN-1998; 98US-0090391.
 PR 03-FEB-1999; 99US-0118570.
 PR 21-JUN-1999; 99US-0337171.
 XX
 PA (CHIR) CHIRON CORP.
 XX
 PI Kennedy GC;
 XX
 PT WPI; 2000-147209/13.
 DR
 XX
 PS Claim 1; Page 38-39; 43pp; English.
 XX
 CC The invention provides polynucleotide sequences (AA258634-648) that are
 CC differentially expressed in pancreatic cancer and dysplasia. The
 CC polynucleotides and polypeptides encoded by them are useful for the
 CC diagnosis and treatment of pancreatic cancer and dysplasia. The
 CC polynucleotides are also useful in forensic, genetic analysis, mapping,
 CC and diagnostic applications if the corresponding region of a gene is
 CC polymorphic in the human population. Therapeutic compositions comprising
 CC antibodies specific for the polypeptides are useful for treating
 CC pancreatic cancer and pancreatic dysplasia as well as other types of

CC cancer e.g. bone cancer, brain tumours, breast cancer, endocrine system
CC cancers e.g. thyroid, pituitary and adrenal glands and the pancreatic
CC islets, gastrointestinal cancers, genitourinary cancers, gynecological
CC cancers, head and neck cancers, leukemia, lymphomas including Hodgkin's
CC and non-Hodgkin's lymphoma, metastatic cancer, myelomas, sarcomas, skin
CC cancer, urinary tract cancers and pediatric cancers. The present
CC sequence represents a polynucleotide sequence that is up-regulated in
CC both dysplasia and pancreatic cancer.
XX
SQ Sequence 539 BP; 165 A; 79 C; 90 G; 205 T; 0 other:

Query Match 27.0%; Score 502; DB 21; Length 539;
Best Local Similarity 96.8%; Pred. No. 2.8e-110;
Matches 522; Conservative 0; Mismatches 16; Indels 1; Gaps 1;

QY 1051 AATAGACATTTACTTCTATGTGGAAGAGTTTTCACAAAGATGCAATTCGTGAA 1110
DB 1 AATAGACATTTACTTCTATGTGGAAGAGTTTTCACAAAGATGCAATTCGTGAA 60
QY 1111 TATTTGTTAGTCCAGCCTATATATTTGCTGTTGGAGTATAGCTGACTCATTTGAAATCA 1170
DB 61 TATTTGTTAGTCCAGCCTATATATTTGCTGTTGGAGTATAGCTGACTCATTTGAAATCA 120
QY 1171 AAGCAATTTTGGAAATTTATGTTTTCATATGCTGTTGTCATGTATAGTTCCTCAG 1230
DB 121 AAGCAATTTTGGAAATTTATGTTTTCATATGCTGTTGTCATGTATAGTTCCTCAG 180
QY 1231 AAACGTGCAATTTGCTTACTTCTATTTTACCTTATGTCATTTATAGCTTAACATACT 1290
DB 181 AAACGTGCAATTTGCTTACTTCTATTTTACCTTATGTCATTTATAGCTTAACATACT 240
QY 1291 CTGCTCCCAATCCAGACTGTTTGTGTAAGTGAAGTCTATGCAATTTGTAATTTGATA 1350
DB 241 CTGCTCCCAATCCAGACTGTTTGTGTAAGTGAAGTCTATGCAATTTGTAATTTGATA 300
QY 1351 ACTTTTTCATCTTTCGAACAAGCTTTTCAGTGGCCCAATTAATGACAGATTCGAAGG 1410
DB 301 ACTTTTTCATCTTTCGAACAAGCTTTTCAGTGGCCCAATTAATGACAGATTCGAAGG 360
QY 1411 TTTATGTTGTAATATGAGTGAATTTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 1469
DB 361 TTTATGTTGTAATATGAGTGAATTTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 420
QY 1470 TTTCAACAAGAACAATGGAATGAGGGAAGAAACATGGAATTTCTTTAGTGCAGTGTGG 1529
DB 421 TTTCAACAAGAACAATGGAATGAGGGAAGAAACATGGAATTTCTTTAGTGCAGTGTGG 480
QY 1530 TCTTCAATTTACATTTAGTTTCTTAAATATTTTAAACATATGTAAGTAATTAAGTG 1588
DB 481 TCTTCAATTTACATTTAGTTTCTTAAATATTTTAAACATATGTAAGTAATTAAGTG 539

RESULT 11
AAK89423
ID AAK89423 standard; DNA; 387 BP.
XX
AC AAK89423;
XX
DT 05-NOV-2001 (first entry)
XX
DE Human digestive system antigen genomic sequence SEQ ID NO: 2999.
XX
RW Human; digestive system antigen; gene therapy; cancer; appendicitis;
KW ulcerative colitis; infection; Hirschsprung's disease; chronic colitis;
KM digestive system disorder; Meckel's diverticulum; ds.
XX
OS Homo sapiens.
XX
PN W0200155314-A2.
XX
PD 02-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US01324.

XX
PR 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226686.
PR 23-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227099.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.

DR WPI: 2001-112729/12.
XX New isolated nucleic acid molecule encoding a SRT polypeptide is useful
PT for production of recombinant SRT polypeptides, gene mapping,
PT diagnosing genetic disorders and for gene therapy -
XX
XX Claim 2; Fig 309; 663pp; English.
XX
CC Sequences AAF93180 - AAF93143 represent polynucleotide sequences encoding
CC human SRT proteins. The cDNA sequences are isolated from various
CC different human tissue cDNA libraries. The invention relates to a method
CC for detecting cDNA encoding an SRT protein, a vector containing cDNA
CC encoding SRT, a host cell transformed with the vector, an isolated SRT
CC polypeptide, and an antibody which binds to SRT. The polynucleotide
CC sequence can be used in gene therapy and is useful in the recombinant
CC production of SRT polypeptides, as a hybridisation probe to screen
CC libraries to isolate cDNAs with sequence identity to SRT polypeptides, to
CC map the gene encoding the SRT polypeptides and analysing genetic
CC disorders, tissue typing and disease tissue detection. The SRT
CC polynucleotide sequences can be used in polymerase chain reaction,
CC screening for new therapeutic molecules and generation of antisense RNA
CC and DNA.
XX
XX Sequence 469 BP; 76 A; 125 C; 119 G; 146 T; 3 other;
Query Match 18.5%; Score 343.2; DB 22; Length 469;
Best Local Similarity 98.6%; Pred. No. 2.8e-72;
Matches 345; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
OY 1 ATGGCGCAGCTAGAGGTTACTGTTTCGCGCCGCTGAGCTGACTGTTTACTGTC 60
DB 120 ATGGCGCAGCTAGAGGTTACTGTTTCGCGCCGCTGAGCTGACTGTTTACTGTC 179
OY 61 TGCCCTCCTCTCTCCGCTTACGCGGCGCTGCGAGAGCCCTACATGAGAGATCTTC 120
DB 180 TGCCCTCCTCTCTCCGCTTACGCGGCGCTGCGAGAGCCCTACATGAGAGATCTTC 239
OY 121 CACCTGCGTCAGGCGCAGCGCTACTGTGAGGGCCATTTCTCCCTTCCAGTGGATCC 180
DB 240 CACCTGCGTCAGGCGCAGCGCTACTGTGAGGGCCATTTCTCCCTTCCAGTGGATCC 299
OY 181 ATGATTTACTACATTAACGCGTGTGACCTGTGACAGTGTGAGAGTGTAAACCTGCCATT 240
DB 300 ATGATTTACTACATTAACGCGTGTGACCTGTGACAGTGTGAGAGTGTAAACCTGCCATT 359
OY 241 TGGATCTTTGGATGTGAGACATGTTGTCTGCTCATTTGGAGTCTCAGATTTGTAAAT 300
DB 360 TGGATCTTTGGATGTGAGACATGTTGTCTGCTCATTTGGAGTCTCAGATTTGTAAAT 419
OY 301 CTTCCTCTCAGTGTGGCACTTCATTTACTATATTTGCTTTCCACAA 350
DB 420 CTTCCTCTCAGTGTGGCACTTCATTTACTATATTTGCTTTCCACAA 469
RESULT 13
AAF93569
ID AAF93569 standard; cDNA; 442 BP.
XX AAF93569;
XX
XX 21-MAY-2001 (first entry)
XX
DE Lung carcinoma cDNA encoding SRT protein SEQ ID 390.
XX
XX Human; SRT; gene therapy; gene mapping; tissue typing; ss.
XX
XX Homo sapiens.
XX
XX WO200107611-A2.
XX
XX 01-FEB-2001.
XX
XX 21-JUL-2000; 2000MO-US20006.
PF

XX 26-JUL-1999; 99US-0145701.
XX
XX (GENE) GENEINTECH INC.
XX
XX Baker KP, Goddard A, Wood WI;
XX
XX WPI: 2001-112729/12.
XX
XX New isolated nucleic acid molecule encoding a SRT polypeptide is useful
PT for production of recombinant SRT polypeptides, gene mapping,
PT diagnosing genetic disorders and for gene therapy -
XX
XX Claim 2; Fig 390; 663pp; English.
XX
CC Sequences AAF93180 - AAF93143 represent polynucleotide sequences encoding
CC human SRT proteins. The cDNA sequences are isolated from various
CC different human tissue cDNA libraries. The invention relates to a method
CC for detecting cDNA encoding an SRT protein, a vector containing cDNA
CC encoding SRT, a host cell transformed with the vector, an isolated SRT
CC polypeptide, and an antibody which binds to SRT. The polynucleotide
CC sequence can be used in gene therapy and is useful in the recombinant
CC production of SRT polypeptides, as a hybridisation probe to screen
CC libraries to isolate cDNAs with sequence identity to SRT polypeptides, to
CC map the gene encoding the SRT polypeptides and analysing genetic
CC disorders, tissue typing and disease tissue detection. The SRT
CC polynucleotide sequences can be used in polymerase chain reaction,
CC screening for new therapeutic molecules and generation of antisense RNA
CC and DNA.
XX
XX Sequence 442 BP; 73 A; 112 C; 115 G; 140 T; 2 other;
Query Match 18.1%; Score 336.4; DB 22; Length 442;
Best Local Similarity 99.1%; Pred. No. 1.2e-70;
Matches 337; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
OY 11 TAGAGGTTACTGTTTCTCGCGCCGCTGAGCTGACTGTTTACTGTCCTCTCT 70
DB 103 TAGAGGTTACTGTTTCTCGCGCCGCTGAGCTGACTGTTTACTGTCCTCTCTCT 162
OY 71 TCTCGGCTTACGCGGCGCTGCGAGAGCCCTACATGAGAGATGTCACCTGCCCTC 130
DB 163 TCTCGGCTTACGCGGCGCTGCGAGAGCCCTACATGAGAGATGTCACCTGCCCTC 222
OY 131 AGGCGACGCTACTGTGAGGCGCCATTTCTCCCTTCCAGTGGAGTCCATGATTA 190
DB 223 AGGCGACGCTACTGTGAGGCGCCATTTCTCCCTTCCAGTGGAGTCCATGATTA 282
OY 191 CATTAACCTGGCTTACTGCTGTGATGATGAGTGTGAAACCTGCCATTTGGATCTTG 250
DB 283 CATTAACCTGGCTTACTGCTGTGATGATGAGTGTGAAACCTGCCATTTGGATCTTG 342
OY 251 GATGATCTGACATGTTGTCTGCTCATTTGGAGTGTGATTTGTTAATCTCTCTCA 310
DB 343 GATGATCTGACATGTTGTCTGCTCATTTGGAGTGTGATTTGTTAATCTCTCTCA 402
OY 311 GTGTTGGCACTTCTATTTACTATATTTGCTTTCCACAA 350
DB 403 GTGTTGGCACTTCTATTTACTATATTTGCTTTCCACAA 442
RESULT 14
AAH98971
ID AAH98971 standard; cDNA; 348 BP.
XX AAH98971;
XX
XX 12-OCT-2001 (first entry)
XX
DE Rat Esr-derived coding sequence SEQ ID NO: 828.
XX
XX Human; sheep; pig; cow; fruit fly; yeast; hamster; macaque; horse;
XX tomato; monkey; dog; sea urchin; expressed sequence tag; EST;
KW

Db 252 AGTTACGGAGGCTTGGGAACGTGAGCT 279

Search completed: September 23, 2003, 11:15:36
Job time : 627 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 23, 2003, 09:16:28 ; Search time 4775 Seconds
(without alignments)
9452.023 Million cell updates/sec

Title: US-10-000-151B-1

Perfect score: 1857

Sequence: 1 atggcgcagctagagggta.....aagtaaatatgncncnaaa 1857

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST.*
1: em_estbda:*
2: em_esthum:*
3: em_estln:*
4: em_estnu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hlc:*
9: gb_est1:*
10: gb_est2:*
11: gb_hlc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rtd:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gss1:*
29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1574	84.8	1833	11	BC022336 Homo sapi
2	872.4	47.0	974	9	AL578231 AL578231
3	760	40.9	958	9	AL555819 AL555819
4	639	34.4	784	12	BI770281 BI770281

5	634.2	34.2	821	12	BI825977
6	618.6	32.3	911	14	CD252003
7	596	32.1	1019	10	BE738701
8	580.4	31.3	1201	9	AL519606
9	527.6	28.4	667	14	CB483386
10	509.4	27.4	667	10	BG613309
11	503.4	27.1	694	12	BI768769
12	469.8	25.3	560	28	A0598053
13	464	25.0	1126	13	BX401726
14	461	24.8	530	28	A0512968
15	459	24.7	724	14	BY740968
16	431	23.2	527	2	HS0086877
17	425	22.9	483	28	A0143712
18	424	22.8	629	14	BY722936
19	423.8	22.8	531	12	BG835696
20	421.6	22.7	659	14	BY734976
21	413.4	22.3	583	13	B0132687
22	401	21.6	535	10	BG692134
23	398.8	21.5	640	10	BB866047
24	391	21.1	468	14	CB158813
25	379.6	20.4	775	9	AJ451344
26	377	20.3	783	13	B0113244
27	374	20.1	816	9	AJ445978
28	368.4	19.8	571	9	AL712776
29	365.8	19.7	862	10	BE738133
30	361.2	19.5	709	13	B0305035
31	360.8	19.4	717	14	BY748369
32	356.8	19.2	491	12	BM149103
33	349	18.8	507	12	BM194113
34	337.8	18.2	721	10	BF031923
35	303	16.3	349	10	BF197055
36	303	16.3	720	13	B0313974
37	296.6	16.0	702	9	AL646668
38	296.4	16.0	623	14	BY745545
39	296.4	16.0	649	14	BY749182
40	294.8	15.9	454	14	CB784843
41	291.8	15.7	500	28	A0118630
42	290.8	15.7	443	12	BM146779
43	272.4	14.7	889	13	B0912625
44	269	14.5	428	9	AA223584
45	268.6	14.5	290	9	AW297936

ALIGNMENTS

RESULT 1
BC022336
LOCUS
DEFINITION Homo sapiens, hypothetical protein FLJ14751, clone IMAGE:4771940,
mRNA.
ACCESSION BC022336
VERSION BC022336.1 GI:18490846
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 1833)
AUTHORS Strausberg, R.
TITLE Direct Submission
JOURNAL Submitted (01-FEB-2002) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA

REMARK
COMMENT NIH-MGC Project URL: <http://mgc.ncl.nih.gov>
Contact: MGC help desk
Email: cgapbs-remail.nih.gov

Tissue Procurement: ATCC
CDNA Library Preparation: Clontech Laboratories, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNI)
DNA Sequencing by: Sequencing group at the Stanford Human Genome
Center, Stanford University School of Medicine, Stanford, CA 94305

Web site: <http://www.shgc.stanford.edu>
 Contact: (Dickson, Mark) medepax11.stanford.edu
 Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers, R. M.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/ILM at: <http://image.llnl.gov>
 Series: IRML Plate: 37 Row: m Column: 24
 This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 14249543
 This clone has the following problem: frame shifted.

FEATURES

source

Location/Qualifiers

1..1833
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /db_xref="IMAGE:4771940"
 /clone="IMAGE:4771940"
 /tissue_type="testis, embryonal carcinoma"
 /lab_host="DH10B"
 /note="Vector: pDNR-LIB"

BASE COUNT 480 a 361 c 344 g 648 t
 ORIGIN

Query Match 84.8%; Score 1574; DB 11; Length 1833;
 Best Local Similarity 97.1%; Pred. No. 5e-237;
 Matches 1620; Conservative 1; Mismatches 46; Indels 2; Gaps 2;

```

OY 1 ATGGGGGAGCTAGAGGTTACTGTTTCGCGCGGCTGAGCTGATACCTTTTACTGTC 60
Db ATGGGCGAGCTGGAAGGTTACTATTCTCGCGGCTGAGCTGATACCTTTTACTGATCC 199
OY 61 TGCCCTCTCTCTCCGCTTCAGCGGCGGCTGCGAGAGCCCTACATGGACGAGATCTTC 120
Db TGCCCTCTCTCTCCGCTTCAGCGGCGGCTGCGAGAGCCCTACATGGACGAGATCTTC 259
OY 121 CACCTGCTCAGCGGCGGCTGAGTGAAGGCCATTTCTCCCTTCCCATGGAGATCCC 180
Db CACCTGCTCAGCGGCGGCTGAGTGAAGGCCATTTCTCCCTTCCCATGGAGATCCC 319
OY 181 ATGATTACATCTACCTGCTGTACCGTGTACGTTGAGTGGAGTGCATACCTCCCAT 240
Db ATGATTACATCTACCTGCTGTACCGTGTACGTTGAGTGGAGTGCATACCTCCCAT 379
OY 241 TGGATCTTGGATGGCTGGAACATGTTGCTGCTCCATTTGGAGTCTCGATTTGTAAT 300
Db TGGATCTTGGATGGCTGGAACATGTTGCTGCTCCATTTGGAGTCTCGATTTGTAAT 439
OY 301 CTTCCTTCAGTGTGGCACTTCTATTATTTGCTTTTCCACAGGTACAAACC 360
Db CTTCCTTCAGTGTGGCACTTCTATTATTTGCTTTTCCACAGGTACAAACC 499
OY 440 CTTCCTTCAGTGTGGCACTTCTATTATTTGCTTTTCCACAGGTACAAACC 499
Db AGAACAAGGCGCTCAAGTATCCAGAGCTGTGCAACATTAACACTACAGATATT 420
OY 500 AGAACAAGGCGCTCAAGTATCCAGAGCTGTGCAACATTAACACTACAGATATT 559
Db 421 CCACACTTATTTTAACTTCTTATTTATACAGAGAGAGATCTATGTTTTTACT 480
OY 560 CCACACTTATTTTAACTTCTTATTTATACAGAGAGAGATCTATGTTTTTACT 619
Db 481 CTTTTCGATTTTGAATGTCTTTATGGAATCATTAACCTTCAGCTTCTGATTT 540
OY 620 CTTTTCGATTTTGAATGTCTTTATGGAATCATTAACCTTCAGCTTCTGATTT 679
Db 541 TGTGCTTCAGTGTGGCAACAATATCATCTGGGCTGTCTCTGTGACAGGAAATGTC 600
OY 660 TGTGCTTCAGTGTGGCAACAATATCATCTGGGCTGTCTCTGTGACAGGAAATGTC 739
Db 601 ATTGACAAAAGTTTAAGGCTTGAAAACTGAGCTACAAAAGAGAGAGACAGACTT 660
OY 740 ATTGACAAAAGTTTAAGGCTTGAAAACTGAGCTACAAAAGAGAGAGACAGACTT 799

```

```

OY 661 CCACCTATTTAAAGGACATTTTCAGCAATTCAGAAAAATCTTCAGTTCTTTGGCTTAT 720
Db 800 CCACCTATTTAAAGGACATTTTCAGCAATTCAGAAAAATCTTCAGTTCTTTGGCTTAT 859
OY 721 TCCATGTCCTTTAAAGGACATTTTCAGCAATTTCTTCAGTTCTTTGGCTTAT 780
Db 860 TCCATGTCCTTTAAAGGACATTTTCAGCAATTTCTTCAGTTCTTTGGCTTAT 919
OY 781 GGAATTCCTTTTGTGCTTTTGTAGTATGTTAAATGTTGAATTTGTTGGCAGT 840
Db 920 GGAATTCCTTTTGTGCTTTTGTAGTATGTTAAATGTTGAATTTGTTGGCAGT 979
OY 841 AGTCATGAAGCTGTCTTCATTTTCTCAACATTTCTACTTTTTCATTTACTCTCTT 900
Db 980 AGTCATGAAGCTGTCTTCATTTTCTCAACATTTCTACTTTTTCATTTACTCTCTT 1039
OY 901 TTTTCTTTTCTCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 960
Db 1040 TTTTCTTTTCTCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1099
OY 961 AAACATGGAATTCG-TTTTGGTGGTACCTTACCTCTCTCTCTCTCTCTCTCTCTCTCT 1019
Db 1100 AAACATGGAATTCG-TTTTGGTGGTACCTTACCTCTCTCTCTCTCTCTCTCTCTCTCT 1159
OY 1020 CACTTATGCTCAATAAATCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1079
Db 1160 CACTTATGCTCAATAAATCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1219
OY 1080 AAGATTTTTCAAAGATATGCAATTCGAAATTAATTTGTTAGTTCAGCCTATATPTTGC 1139
Db 1220 AAGATTTTTCAAAGATATGCAATTCGAAATTAATTTGTTAGTTCAGCCTATATPTTGC 1279
OY 1140 TGGTGGAGATATAGCTGACATCAATCAAAATCAAAATCAAAATCAAAATCAAAATCA 1199
Db 1280 TGGTGGAGATATAGCTGACATCAATCAAAATCAAAATCAAAATCAAAATCAAAATCA 1339
OY 1200 CATATGCTTTGTCATTTGTTATAGTTCCTCAGAAATCTCTGGAATTTCTTACTTATTT 1259
Db 1340 CATATGCTTTGTCATTTGTTATAGTTCCTCAGAAATCTCTGGAATTTCTTACTTATTT 1399
OY 1260 ACCATTATGCTTTATAGCTTATACATTAACCTGCTCCCATCAGACTTCTTGTGGA 1319
Db 1400 ACCATTATGCTTTATAGCTTATACATTAACCTGCTCCCATCAGACTTCTTGTGGA 1459
OY 1320 ACTGAGTTCATATGCAATTTGTTAATTTTCAATCTTTTACATCTTTTGAACAAGACTT 1379
Db 1460 ACTGAGTTCATATGCAATTTGTTAATTTTCAATCTTTTACATCTTTTGAACAAGACTT 1519
OY 1380 TCAGTGGCCAAATAGTCAAGCAATTCAAAGTTTATGTTGTTATATCACTGATTTTGA 1439
Db 1520 TCAGTGGCCAAATAGTCAAGCAATTCAAAGTTTATGTTGTTATATCACTGATTTTGA 1579
OY 1440 ACTGTAATAATGAGCTTTAATA-TAGACCATTTCTACAAAGAACAGAAAGAGNGGA 1498
Db 1580 ACTGTAATAATGAGCTTTAATA-TAGACCATTTCTACAAAGAACAGAAAGAGNGGA 1639
OY 1499 AACATGGAATTTCTTTTGGTGCAGCTGTGCTCTCAATATACATTAAGTTTTTAAAT 1558
Db 1640 AACATGGAATTTCTTTTGGTGCAGCTGTGCTCTCAATATACATTAAGTTTTTAAAT 1699
OY 1559 ATATTTTAAATATATGTAAGAAATTAAGTGGCAAGAACTGGAAGCTTAAAGCTGCT 1618
Db 1700 ATATTTTAAATATATGTAAGAAATTAAGTGGCAAGAACTGGAAGCTTAAAGCTGCT 1759
OY 1619 TCAAAAGCCTGAAATATGGAATAAANMNGTTTNCAGATATTCATTA 1667
Db 1760 TCAAAAGCCTGAAATAATGGAATAAATATTTGTTTCAGATATTCACAAA 1808

```

RESULT 2
 LOCUS AL578231/c 974 bp mRNA linear EST 01-JUN-2003
 DEFINITION AL578231 Homo sapiens HELA CELLS COT 25-NORMALIZED Homo sapiens

CDNA clone CS0DK002YI20 3-PRIME, mRNA sequence.
 AL578231
 VERSION AL578231.2 GI:31316449
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 1 (bases 1 to 974)
 L.I.W.B., Gruber,C., Jessee,J. and Polayes,D.
 Full-length cDNA libraries and normalization
 Unpublished
 On Feb 10, 2001 this sequence version replaced gi:12942110.
 CONTACT: Genoscope
 Genoscope - Centre National de Sequencage
 BP 191 91006 EVRY cedex - France
 Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
 Library was constructed by Life Technologies, a division of
 Invitrogen. This sequence belongs to sequence cluster 6664.f for
 more information about this cluster, see
 http://www.genoscope.cns.fr/
 cgi-bin/cluster.cgi?seq=CS0DK002BE10NP1&cluster=6664.f. Contact :
 Peng Liang Email : filiang@life.techn.com URL :
 http://fulllength.invitrogen.com/Invitrogen Corporation 1600
 Faraday Avenue Genoscope sequence ID : CS0DK002BE10NP1.
 Location/Qualifiers
 1..974
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CS0DK002YI20"
 /cell_type="HELA CELLS COT 25-NORMALIZED"
 /cell_line="HELA"
 /clone_lib="Homo sapiens HELA CELLS COT 25-NORMALIZED"
 /note="1st strand cDNA was primed with a NotI-oligo(dT)
 primer. Five prime end enriched, double-strand cDNA was
 digested with Not I and cloned into the Not I and EcoR V
 sites of the pCMVSPORT 6 vector. Library was normalized."
 BASE COUNT 344 a 163 c 170 g 263 t 34 others
 ORIGIN
 Query Match 47.0%; Score 872.4; DB 9; Length 974;
 Best Local Similarity 91.8%; Pred. No. 4,7e-127;
 Matches 893; Conservative 29; Mismatches 50; Indels 1; Gaps 1;
 682 GCAGATTGCAAGAAATTTCTGAGTTCTTTGGCTATTCAGTCTTAATAACTTG 741
 |||||||
 974 GCAGATTGCAAGAAATTTCTGAGTTCTTTGGCTATTCAGTCTTAATAACTTG 915
 |||||||
 742 AGTATGCTTTCTGTTGACCTGGCCCTACATCCTTCGGGATTTGTTGCTTTT 801
 |||||||
 914 AGTATGCTTTGCTTGACCTGGCCCTACACCTTCGGGATTTGTTGCTTTT 855
 |||||||
 802 GTAGTACTTAATGCTGGAATTTGTTATGGCGATCGAGTAGTCATGAAGCCTGCTTCAT 861
 |||||||
 854 GTAGTACTTAATGCTGGAATTTGTTATGGCGATCGAGTAGTCATGAAGCCTGCTTCAT 795
 |||||||
 862 TTTCCTCAACTATTTCTACTTTTTCATACCTCTTTTCTTTCTTCCCACTGCTG 921
 |||||||
 794 TTTCCTCAACTATTTCTACTTTTTCATACCTCTTTTCTTTTCCCACTGCTG 735
 |||||||
 922 TCTCCTAGCAAAATTAAGACTTTTCTTCCCTAGTTTGAACAATGAAATTCCTTTTGG 981
 |||||||
 734 TCTCCTAGCAAAATTAAGACTTTTCTTCCCTAGTTTGAACAATGAAATTCCTTTTGG 675
 |||||||
 982 GTGGTTACCTTAGTCTGCTGTTTGAATTTGAATTCATATGCTCATTAATACTTG 1041
 |||||||
 674 GTGGTTACCTTAGTCTGCTGTTTGAATTTGAATTCATATGCTCATTAATACTTG 615
 |||||||
 1042 CTAGCAGACAAATAGACTTTACTTTCTATGCTGGGAAAGAGTTTTCAAAGATATGCA 1101
 |||||||
 614 CTAGCAGACAAATAGACTTTACTTTCTATGCTGGGAAAGAGTTTTCAAAGATATGCA 555
 |||||||

1102 ATTCTGAATATTTAGTTAGTTCAGCCTATATATTTGCTGGTGGAGTATAGCTGACTCA 1161
 |||||||
 554 ACAGAAAATATTTAGTTCAGCCTATATATTTGGTGGAGTATAGCTGACTCA 495
 |||||||
 1162 TTGAATTCAGCAATTTTGGCAATTTAATGTTTTCATATGCTGTTGTTATTA 1221
 |||||||
 494 TTGAATTCAGCAATTTTGGCAATTTAATGTTTTCATATGCTGTTGTTATTA 435
 |||||||
 1222 GTTCTCAGAAAGCTGGAATTTGTTTACTTCAATTTTACTTATGATATTAAGGCTT 1281
 |||||||
 434 GTTCTCAGAAAGCTGGAATTTGTTTACTTCAATTTTACTTATGATATTAAGGCTT 375
 |||||||
 1282 AACATTAATCTGCTCCACATCCAGACTGTTGTTGTAAGTGAATGCTATGCAATGTT 1341
 |||||||
 374 AACATTAATCTGCTCCACATCCAGACTGTTGTTGTAAGTGAATGCTATGCAATGTT 315
 |||||||
 1342 AATTCTAATCTTTTACATCTTTCTGTAACAAGACTTTTCAAGTGGCCAAATAGTCAGAC 1401
 |||||||
 314 AATTCTAATCTTTTACATCTTTCTGTAACAAGACTTTTCAAGTGGCCAAATAGTCAGAC 255
 |||||||
 1402 ATTCAAGGTTTATGTTGTAATATCATGATATTTTGAAGTGAATGAGCTTAATTA 1461
 |||||||
 254 ATTCAAGGTTTATGTTGTAATATCATGATATTTTGAAGTGAATGAGCTTAATTA 195
 |||||||
 1462 -TAGACCATTTCTCAAGAAACAAGTGAATGAGGAAACATGGAATTTCTTTAGGTG 1520
 |||||||
 194 TTACACCATTTCTCAAGAAACAAGTGAATGAGGAAACATGGAATTTCTTTAGGTG 135
 |||||||
 1521 CAGTGTGCTGCTTCAATTTACATTTAGTTTATATATATTTTAAACATATGTAAGAA 1580
 |||||||
 134 CAGTGTGCTGCTTCAATTTACATTTAGTTTATATATATTTTAAACATATGTAAGAA 75
 |||||||
 1581 ATTAAGTGGCAAGAAAGCTGGAAGCTTAAGACCTGCTCAAGAGCTGAATTAATGGA 1640
 |||||||
 74 ATTAAGTGGCAAGAAAGCTGGAAGCTTAAGACCTGCTCAAGAGCTGAATTAATGGA 15
 |||||||
 1641 AATTAANWNGTTT 1653
 |||||||
 14 AVDAAGKGGTNT 2
 |||||||
 RESULT 3
 AL555819 958 bp mRNA linear EST 31-MAY-2003
 LOCUS
 DEFINITION
 AL555819 Homo sapiens HELA CELLS COT 25-NORMALIZED Homo sapiens
 ACCESSION
 AL555819
 VERSION
 AL555819.2 GI:31277625
 KEYWORDS
 EST.
 SOURCE
 Homo sapiens (human)
 ORGANISM
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 1 (bases 1 to 958)
 L.I.W.B., Gruber,C., Jessee,J. and Polayes,D.
 Full-length cDNA libraries and normalization
 Unpublished
 On Feb 15, 2001 this sequence version replaced gi:12897913.
 CONTACT: Genoscope
 Genoscope - Centre National de Sequencage
 BP 191 91006 EVRY cedex - France
 Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
 Library was constructed by Life Technologies, a division of
 Invitrogen. This sequence belongs to sequence cluster 6664.f for
 more information about this cluster, see
 http://www.genoscope.cns.fr/
 cgi-bin/cluster.cgi?seq=CS0DK002BE10NP1&cluster=6664.f. Contact :
 Peng Liang Email : filiang@life.techn.com URL :
 http://fulllength.invitrogen.com/Invitrogen Corporation 1600
 Faraday Avenue Genoscope sequence ID : CS0DK002BE10NP1.
 Location/Qualifiers
 1..958
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CS0DK002YI20 5-PRIME, mRNA sequence."
 /cell_type="HELA CELLS COT 25-NORMALIZED"
 /cell_line="HELA"
 /clone_lib="Homo sapiens HELA CELLS COT 25-NORMALIZED"
 /note="1st strand cDNA was primed with a NotI-oligo(dT)
 primer. Five prime end enriched, double-strand cDNA was
 digested with Not I and cloned into the Not I and EcoR V
 sites of the pCMVSPORT 6 vector. Library was normalized."
 BASE COUNT 344 a 163 c 170 g 263 t 34 others
 ORIGIN
 Query Match 47.0%; Score 872.4; DB 9; Length 974;
 Best Local Similarity 91.8%; Pred. No. 4,7e-127;
 Matches 893; Conservative 29; Mismatches 50; Indels 1; Gaps 1;
 682 GCAGATTGCAAGAAATTTCTGAGTTCTTTGGCTATTCAGTCTTAATAACTTG 741
 |||||||
 974 GCAGATTGCAAGAAATTTCTGAGTTCTTTGGCTATTCAGTCTTAATAACTTG 915
 |||||||
 742 AGTATGCTTTCTGTTGACCTGGCCCTACATCCTTCGGGATTTGTTGCTTTT 801
 |||||||
 914 AGTATGCTTTGCTTGACCTGGCCCTACACCTTCGGGATTTGTTGCTTTT 855
 |||||||
 802 GTAGTACTTAATGCTGGAATTTGTTATGGCGATCGAGTAGTCATGAAGCCTGCTTCAT 861
 |||||||
 854 GTAGTACTTAATGCTGGAATTTGTTATGGCGATCGAGTAGTCATGAAGCCTGCTTCAT 795
 |||||||
 862 TTTCCTCAACTATTTCTACTTTTTCATACCTCTTTTCTTTCTTCCCACTGCTG 921
 |||||||
 794 TTTCCTCAACTATTTCTACTTTTTCATACCTCTTTTCTTTTCCCACTGCTG 735
 |||||||
 922 TCTCCTAGCAAAATTAAGACTTTTCTTCCCTAGTTTGAACAATGAAATTCCTTTTGG 981
 |||||||
 734 TCTCCTAGCAAAATTAAGACTTTTCTTCCCTAGTTTGAACAATGAAATTCCTTTTGG 675
 |||||||
 982 GTGGTTACCTTAGTCTGCTGTTTGAATTTGAATTCATATGCTCATTAATACTTG 1041
 |||||||
 674 GTGGTTACCTTAGTCTGCTGTTTGAATTTGAATTCATATGCTCATTAATACTTG 615
 |||||||
 1042 CTAGCAGACAAATAGACTTTACTTTCTATGCTGGGAAAGAGTTTTCAAAGATATGCA 1101
 |||||||
 614 CTAGCAGACAAATAGACTTTACTTTCTATGCTGGGAAAGAGTTTTCAAAGATATGCA 555
 |||||||

```

/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CSODK002Y120"
/cell_line="HELA CELLS COT 25-NORMALIZED"
/clone_lib="HELA"
/clone_lib="Homo sapiens HELA CELLS COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."
BASE COUNT      221 a      213 c      206 g      310 t      8 others
ORIGIN

```

```

Query Match      40.9%; Score 760; DB 9; Length 958;
Best Local Similarity 97.5%; Pred. No. 1.9e-109;
Matches 780; Conservative 2; Mismatches 17; Indels 1; Gaps 1;

QY 10 CTAGAGGTTACTGTTCTTCGCGCCCTTGAGCTTACTTTTGTGCTGCTGCTC 69
DB 160 STGGAAGGTACTATTCTCGCGCCCTTGAGCTTACTTTTGTGCTGCTGCTC 219
QY 70 TTCTCCGCTTCAGCGCGCGCTGAGAGCCCTACATGAGAGATCTCCACCTGCT 129
DB 220 TTCTCCGCTTCAGCGCGCGCTGAGAGCCCTACATGAGAGATCTCCACCTGCT 279
QY 130 CAGGCGACGCTACTGTAGAGGCCATTCTCCCTTCCAGTGGATCCATGATTA 189
DB 280 CAGGCGCA-SGCTACTGTAGAGGCCATTCTCCCTTCCAGTGGATCCATGATTA 338
QY 190 ACATTACTGCTTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 249
DB 339 ACATTACTGCTTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 398
QY 250 GGATGCTGCAACATGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 309
DB 399 GGATGCTGCAACATGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 458
QY 310 AGTGTGCAACTTCTATTACTATTATTGCTTTTCCACAGATCAACCCAGAAACA 369
DB 459 AGTGTGCAACTTCTATTACTATTATTGCTTTTCCAGAGTCAACCCAGAAACA 518
QY 370 GCTGCTCAAGTATCAGAGGCTGTCACATTAACATGAGTATTTCCACACTT 429
DB 519 GCTGCTCAAGTATCAGAGGCTGTCACATTAACATGAGTATTTCCACACTT 578
QY 430 TATTTTAACTTCTTATTATATACAGAGGATCTATGTTTAACTCTTTTGA 489
DB 579 TATTTTAACTTCTTATTATATACAGAGGATCTATGTTTAACTCTTTTGA 638
QY 490 TATTTGATGTCCTTTATGGAATCATAAACTTACGCTTCTTGATTTTGTGCT 549
DB 639 TATTTGATGTCCTTTATGGAATCATAAACTTACGCTTCTTGATTTTGTGCT 698
QY 550 ATGTTGGCAAAACAATATCATCTGGGCTGCTGCTGCTGCTGCTGCTGCTGCT 609
DB 699 ATGTTGGCAAAACAATATCATCTGGGCTGCTGCTGCTGCTGCTGCTGCTGCT 758
QY 610 AAGTTAACTGAGGCTTGGAAAACTGAGTACAAAGAAGAGAGACTTCCACTAT 669
DB 759 AAGTTAACTGAGGCTTGGAAAACTGAGTACAAAGAAGAGAGACTTCCACTAT 818
QY 670 AAAGGACATTTGCAAAATTCAGAAAAATCTTCAGTTTCTTTGGCTTATTCAT 729
DB 819 AAAGGACATTTGCAAAATTCAGAAAAATCTTCAGTTTCTTTGGCTTATTCAT 878
QY 730 TTTAAAACTGAGATGCTTTTCTGTTGACTTGGCCCTACATCTCTGGGATTTCT 789
DB 879 TTTAAAACTGAGATGCTTTTCTGTTGACTTGGCCCTACATCTCTGGGATTTCT 938
QY 790 TTTTGTGCTTTGTAGTAGT 809
DB 939 TTTTGTGCTTTGTAGTAGT 958

```

```

RESULT 4
LOCUS      B1770281
DEFINITION 603055973P1 NIH_MGC_122 Homo sapiens cDNA clone IMAGE:5205402 5',
            mRNA sequence.
ACCESSION  B1770281
VERSION    B1770281.1 GI:15761859
KEYWORDS  EST.
SOURCE     Homo sapiens (human)
ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE  1 (bases 1 to 784)
            NIH-MGC http://mgi.nci.nih.gov/.
            National Institutes of Health, Mammalian Gene Collection (MGC)
            JOURNAL
            Unpublished
            Contact: Robert Strausberg, Ph.D.
            Email: cgaabs-remail.nih.gov
            Tissue Procurement: Life Technologies, Inc.
            cDNA Library Preparation: Life Technologies, Inc.
            DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM1515 row: a column: 19
            High quality sequence stop: 756.
            Location/Qualifiers
                source
                    1..784
                        /organism="Homo sapiens"
                        /mol_type="mRNA"
                        /db_xref="taxon:9606"
                        /clone="IMAGE:5205402"
                        /lab_host="DH10B"
                        /clone_lib="NIH_MGC_122"
                        /note="Organ: pooled lung and spleen; Vector: pCMV-Sport6;
                        Site_1: NotI; Site_2: EcoRV (destroyed); RNA source
                        anonymous pool of 24 week female lung, 16 week female
                        spleen, and 20-22 week male spleens. Library is oligo-dT
                        primed and directionally cloned (EcoRV site is destroyed
                        upon cloning). Average insert size 1.4 kb, insert size
                        range 1-3 kb. Library is normalized and enriched for
                        full-length clones and was constructed by C. Gruber
                        (Invitrogen). Research Genetics tracking code 026. Note:
                        this is a NIH-MGC Library."
BASE COUNT      180 a      187 c      170 g      247 t
ORIGIN
Query Match      34.4%; Score 639; DB 12; Length 784;
Best Local Similarity 97.6%; Pred. No. 1.8e-90;
Matches 659; Conservative 0; Mismatches 15; Indels 1; Gaps 1;

QY 1 ATGGGCGAGCTAGAGGTTACTGTTTCGCGCGCTGAGCTGATCTTTTACTGTC 60
DB 87 ATGGGCGAGCTAGAGGTTACTGTTTCGCGCGCTGAGCTGATCTTTTACTGTC 146
QY 61 TGCCCTCTCTCTCCGCTTACGAGCGGCGCTGCGAGAGCCCTACATGAGAGATCT 120
DB 147 TGCCCTCTCTCTCCGCTTACGAGCGGCGCTGCGAGAGCCCTACATGAGAGATCT 206
QY 121 CACCTGCTCAGGCGCAGCGCTACTGTGAGGGCCATTCTCCCTTCCAGTGGGATCC 180
DB 207 CACCTGCTCAGGCGCAGCGCTACTGTGAGGGCCATTCTCCCTTCCAGTGGGATCC 266
QY 181 ATGATTACTACTATTACCTGCGCTTGTACTGGTGTAGTGTGAGTGTCAAACTCCATT 240
DB 267 ATGATTACTACTATTACCTGCGCTTGTACTGGTGTAGTGTGAGTGTCAAACTCCATT 326
QY 241 TGGATCTTGGATGTGTGACATGTTGCTGCTCATTTGGATGCTCAGATTGTAT 300
DB 327 TGGATCTTGGATGTGTGACATGTTGCTGCTCATTTGGATGCTCAGATTGTAT 386

```

QY 301 CTTCTCTCAGTGTGGCAACTCTCTATTAATTAATTTGCTTTCCAAAGGTACAACCC 360
 Db 387 CTTCTCTCAGTGTGGCAACTCTCTATTAATTAATTTGCTTTCCAAAGGTACAACCC 446
 QY 361 AGAACAAGGCTGCTCAACTATCCAGAGAGTCTTGCAACATTACACATGACATATT 420
 Db 447 AGAACAAGGCTGCTCAACTATCCAGAGAGTCTTGCAACATTACACATGACATATT 506
 QY 421 CCAACATTTATTTTAACTCTCTATTAATTAATTAATTTGCTTTCCAAAGGTACAACCC 480
 Db 507 CCAACATTTATTTTAACTCTCTATTAATTAATTAATTTGCTTTCCAAAGGTACAACCC 566
 QY 481 CTTTTCATATTTTGAATGATGCTTTTGAATATCAATAAATTCAGCTTCTTGAATTT 540
 Db 567 CTTTTCATATTTTGAATGATGCTTTTGAATATCAATAAATTCAGCTTCTTGAATTT 626
 QY 541 TGTGGCTTCATGTTTGGCAACAATATATCTGGGCTCTTCTTGTGAGGA-ATGT 599
 Db 627 CGTGGCTTCATGTTTGGCAACAATATATCTGGGCTCTTCTTGTGAGGA-ATGT 686
 QY 600 CATTCGCAAAAGTTAACTGAGCTTGGAAGACTGAGCTCAAAAGAGAGACAGACT 659
 Db 687 CATTCGCAAAAGTTAACTGAGCTTGGAAGACTGAGCTCAAAAGAGAGACAGACT 746
 QY 660 TCACCTATTAAAG 674
 Db 747 TCACCTATTAAAG 761

RESULT 5
 B1825977 821 bp mRNA linear EST 04-OCT-2001
 LOCUS 60307650F1 NIH_MGC_119 Homo sapiens cDNA clone IMAGE:5168319 5',
 DEFINITION mRNA sequence.

ACCESSION B1825977
 VERSION B1825977.1 GI:15937527
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 821)
 AUTHORS NIH-MGC http://mgi.nci.nih.gov/
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgabs-remail.nih.gov
 Tissue Procurement: Life Technologies, Inc.
 cDNA Library Preparation: Life Technologies, Inc.
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LINL at:
 http://image.linl.gov
 Plate: L14M1418 row: h column: 16
 High quality sequence start: 6
 High quality sequence stop: 800.

FEATURES
 source
 location/Qualifiers
 1..821

/organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:5168319"
 /tissue_type="medulla"
 /lab_host="DH10B"
 /clone_lib="NIH_MGC_119"
 /note="Organ: brain; Vector: pCMV-SPORT6; Site_1: NotI;
 Site_2: EcoRV (destroyed); RNA source normal medulla from
 anonymous male age 27. Library is oligo-dT primed and
 directionally cloned (EcoRV site is destroyed upon
 cloning). Average insert size 1.3 kb, insert size range
 0.9-3 kb. Library is normalized and enriched for
 full-length clones and was constructed by C. Gruber

(Invertegen). Research Genetics tracking code 013. Note:
 this is a NIH_MGC library.
 BASE COUNT 187 a 195 c 179 g 260 t
 ORIGIN

Query Match 34.2%; Score 634.2; DB 12; Length 821;
 Best Local Similarity 97.4%; Pred. No. 1e-89;
 Matches 708; Conservative 0; Mismatches 13; Indels 6; Gaps 6;

QY 1 ATGGCGAGCTAAGGTTACTGTCTGCGGCCCTTGAGCTGTACTTTTAGTGTCC 60
 Db 101 ATGGCGAGCTAAGGTTACTGTCTGCGGCCCTTGAGCTGTACTTTTAGTGTCC 160
 QY 61 TGCCTCTCTTCCCGCTTCAGCGGCGCTGAGAGCCCTACATGAGAGATCTTC 120
 Db 161 TGCCTCTCTTCCCGCTTCAGCGGCGCTGAGAGCCCTACATGAGAGATCTTC 219
 QY 121 CACCTGCTGAGCGCAGCGCTACTGTGAGGGCCATTCTCCCTTCCAGTGGATCCC 180
 Db 220 CACCTGCTGAGCGCAGCGCTACTGTGAGGGCCATTCTCCCTTCCAGTGGATCCC 279
 QY 181 ATGATTACTATTAATCTGCTGTACTGTGTCACTGTGAGTGTGCAACCTGCAATT 240
 Db 280 ATGATTACTATTAATCTGCTGTACTGTGTCACTGTGAGTGTGCAACCTGCAATT 339
 QY 241 TGGATCTTTGATGTGCTGAACATGTTGCTGCCATTGGGATGCTAGATTGTTAAT 300
 Db 340 TGGATCTTTGATGTGCTGAACATGTTGCTGCCATTGGGATGCTAGATTGTTAAT 397
 QY 301 CTTCTCTCAGTGTGGCAACTCTCTATTAATTAATTTGCTTTCCAAAGGTACAACCC 360
 Db 398 CTTCTCTCAGTGTGGCAACTCTCTATTAATTAATTTGCTTTCCAAAGGTACAACCC 457
 QY 361 AGAACAAGGCTGCTCAACTATCCAGAGAGTCTTGCAACATTACACATGACATATT 420
 Db 458 AGAACAAGGCTGCTCAACTATCCAGAGAGTCTTGCAACATTACACATGACATATT 517
 QY 421 CCAACATTTATTTTAACTCTCTATTAATTAATTAATTTGCTTTCCAAAGGTACAACCC 480
 Db 518 CCAACATTTATTTTAACTCTCTATTAATTAATTAATTTGCTTTCCAAAGGTACAACCC 577
 QY 481 CTTTTCATATTTTGAATGATGCTTTTGAATATCAATAAATTCAGCTTCTTGAATTT 540
 Db 578 CTTTTCATATTTTGAATGATGCTTTTGAATATCAATAAATTCAGCTTCTTGAATTT 635
 QY 541 TGTGGCTTCATGTTTGGCAACAATATATCTGGGCTCTTCTTGTGAGGA-ATGT 599
 Db 636 TGTGGCTTCATGTTTGGCAACAATATATCTGGGCTCTTCTTGTGAGGA-ATGT 695
 QY 601 ATTGCACAAAGTTAACTGAGCTTGGAAGACTGAGCTCAAAAGAGAGACAGACTT 660
 Db 696 ATTGCACAAAGTTAACTGAGCTTGGAAGACTGAGCTCAAAAGAGAGACAGACTT 755
 QY 661 CCACTATTAAAGGACCATTTGAGAAATTCAGAAATTTCTTTCAGTTTGGCTTAT 720
 Db 756 CCACTATTAAAGGACCATTTGAGAAATTCAGAAATTTCTTTCAGTTTGGCTTAT 814
 QY 721 TCCATGT 727
 Db 815 TCCATGT 821

RESULT 6
 CD252003 911 bp mRNA linear EST 22-MAY-2003
 LOCUS AGENCOURT_14211757 NIH_MGC_180 Homo sapiens cDNA clone
 DEFINITION IMAGE:30383501 5', mRNA sequence.
 ACCESSION CD252003
 VERSION CD252003.1 GI:31012469
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE 1 (bases 1 to 911)
 AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished
 COMMENT Contact: Daniela S. Gerhard, Ph.D.
 Office of Cancer Genomics
 National Cancer Institute / NIH
 Bldg. 31 Rm10A07 Bethesda, MD 20892
 Email: cgaabs-remail.nih.gov
 Tissue Procurement: Dr. Michael Brownstein
 CDNA Library Preparation: Invitrogen Corp
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNLN at:
<http://image.lnl.gov>
 Plate: NDAM47 row: h column: 06
 High quality sequence stop: 665.
 Location/Qualifiers
 1..911

FEATURES
 source /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone IMAGE:30383501"
 /lab_host="DH10B-Ton A (T1 and T5 phage resistances)"
 /clone_lib="NIH_MGC_180"
 /note="Organ: Testis; Vector: PCMV-SPORT.1; Site.1: NotI; Site.2: EcoRV (destroyed); Library is oligo-dr primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.68 kb. Library was constructed by (Invitrogen). Note: this is a NIH-MGC Library."
 location/Qualifiers
 1..911

BASE COUNT 196 a 211 c 199 g 305 t

Query Match 33.3%; Score 618.6; DB 14; Length 911;
 Best Local Similarity 90.2%; Pred. NO. 2.7e-87;
 Matches 708; Conservative 0; Mismatches 44; Indels 33; Gaps 3;

OY 1 ATGGCGCAGTAAAGAGGTTACTGTTTCGGCGCCCTGAGCTGACCTTTTAACTGTC 60
 DB 128 ATGGCGCAGTAAAGAGGTTACTGTTTCGGCGCCCTGAGCTGACCTTTTAACTGTC 187
 OY 61 TGCCCTCTCTTCGCGCTTACAGCGGCGCTGCGAGAGCCCTACAGTGAAGATCTTC 120
 DB 188 TGCCCTCTCTTCGCGCTTACAGCGGCGCTGCGAGAGCCCTACAGTGAAGATCTTC 247
 OY 121 CACCTGCCTCAGGCGCAGCTACTGTCGAGGCCATTTCCTTCCAGTGGATGCC 180
 DB 248 CACCTGCCTCAGGCGCAGCTACTGTCGAGGCCATTTCCTTCCAGTGGATGCC 281
 OY 181 ATGATTACTAATTACCTGCGCTTGTACCTGGTGCATTTGGAGTGAACCTGCATT 240
 DB 282 ATGATTACTAATTACCTGCGCTTGTACCTGGTGCATTTGGAGTGAACCTGCATT 341
 OY 241 TGGATCTTGGATGGTGTGAACATGTTGTCTCCATTTGGAGTGAACCTGCATT 300
 DB 342 TGGATCTTGGATGGTGTGAACATGTTGTCTCCATTTGGAGTGAACCTGCATT 401
 OY 301 CTTCCTCTCACTGTTGGCACTTGTATTTACTATTTGCTTTTCCAAAGGTACAAACC 360
 DB 402 CTTCCTCTCACTGTTGGCACTTGTATTTACTATTTGCTTTTCCAAAGGTACAAACC 461
 OY 361 AGAACAAGGCTGCCTCAAGTATCCAGAGAGTCTTGCATATTAACTAGACATATTT 420
 DB 462 AGAACAAGGCTGCCTCAAGTATCCAGAGAGTCTTGTAACTATTAAGCTAGCAATTT 521
 OY 421 CCAACACTTATTTTAACTTCTTATTTATACAGAGAGATCTATGTTTAACT 480
 DB 522 CCAACACTTATTTTAACTTCTTATTTATACAGAGAGATCTATGTTTAACT 581
 OY 481 CTTTTCATATTTGATGTCTCTTATGGAATCATAAACTTCAGCTTCCCTGGATTT 540
 DB 582 CTTTTCATATTTGATGTCTCTTATGGAATCATAAACTTCAGCTTCCCTGGATTT 641

DB 582 CTTTTCATATTTGATGTCTCTTATGGAATCATAAACTTCAGCTTCCCTGGATTT 641
 OY 541 TGTGGCTTCATGTTTCGGCAACAATATCATATTCATGCGCTGCTCTGTGACGGAATGTC 600
 DB 642 TGTGGCTTCATGTTTCGGCAACAATATCATATTCATGCGCTGCTCTGTGACGGAATGTC 701
 OY 601 ATTGCACAAAGTTAACTGAGGCTTGGCAAACTGAGCTACAAAGAAGAGACAGACTT 660
 DB 702 ATTGCACAAAGTTAACTGAGGCTTGGCAAACTGAGCTACAAAGAAGAGACAGACTT 761
 OY 661 --CCACCATTAAGAGACCATTTGCAGATTCAGAAAA-----TTCCTGAGTTCTTTT 713
 DB 762 TCCACCTATTTAAAGACCATTTGCAGATTCAGAAAAATTCCTGAGTTCTTTT 821
 OY 714 GGCCTTATTCATGCTCTTTAAACTGAGTATGCTTTCTGTTGACTTGGCCCTACAT 773
 DB 822 GCTTATTTCCATGTCCTTTAAACTTGAATGCTTTTCTGTTGACTTGGCCCTACAT 881
 OY 774 CCTTC 778
 DB 882 CCTAC 886

RESULT 7
 BE738701/c 1019 bp mRNA linear EST 15-SEP-2000
 LOCUS 6015729441 NIH_MGC_57 Homo sapiens CDNA clone IMAGE:3839674 3',
 DEFINITION mRNA sequence.
 BE738701
 ACCESSION BE738701.1 GI:10152693
 VERSION EST.
 KEYWORDS Homo sapiens (human)
 SOURCE Homo sapiens
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 REFERENCE 1 (bases 1 to 1019)
 AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgaabs-remail.nih.gov
 Tissue Procurement: ATCC
 CDNA Library Preparation: CLONTECH Laboratories, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNLN at:
<http://image.lnl.gov>.
 Plate: LDCM528 row: h column: 11
 High quality sequence start: 7
 High quality sequence stop: 721.
 Location/Qualifiers
 1..1019

FEATURES
 source /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone IMAGE:3839674"
 /tissue_type="g1oblastoma"
 /lab_host="DH10B (T1 phage-resistant)"
 /clone_lib="NIH_MGC_57"
 /note="Organ: brain; Vector: pDNR-LIB (Clontech); Site.1: SfiI (99cgcctcggcc); Site.2: SfiI (99ccatctggcc); Double-stranded cDNA was prepared from cell line RNA. 5' and 3' adaptors were used in cloning as follows: 5' adaptor sequence: 5'-ATCTAGAGGCCAGAGGCGCCGACATG-dt(30)BN-3' (where B = A, C, or G and N = A, C, G, or T). Average insert size 1.55 kb (range 0.9-4.0 kb). 12/15 colonies contained inserts by PCR. This library was enriched for full-length clones and was constructed by Clontech Laboratories (Palo Alto, CA)."
 BASE COUNT 395 a 168 c 177 g 279 t

Query Match	32.1%;	Score 596;	DB 10;	Length 1019;
Best Local Similarity	89.5%;	Pred. No. 9.4e-84;		
Matches 719;	Conservative 0;	Mismatches 76;	Indels 8;	Gaps 7;
QY	798	TTTTGAGTAGTTAATGCTGAATTTTATGCGCATGCGAGTAGTCATGAACCTGCTCT	857	
DB	802	TGCTCTTGTATATATATGCTGATGCTGATGCAATCTGACTCCATGCAAGTCCGTC	743	
QY	858	TCATTTCTCCATCTACTCTTTTTCATTTCTCTCTTTTCTCTTCTCTCTCTCTCT	917	
DB	742	CTAATCTTCTCACTATTAACCTGTTTTCATTTCTACTCTCTATTTTCCTCTCTCT	684	
QY	918	CCTGCTCTCCATGCAAAATTAAGACTTTTCTCTCTCTCTCTCTCTCTCTCTCT	976	
DB	683	CCGTCTCTCCATGCAAAATTAAGACTTTTCTCTCTCTCTCTCTCTCTCTCTCT	624	
QY	977	TTTTGGTGTACCTTACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	1034	
DB	623	TTTTTGTGTACCTTACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	564	
QY	1035	ATACCTGCTAGCAGCAATAGACA--TTATCTTCTATGCTGGAAGAGTTTTCACA	1093	
DB	563	ATACCTGCTAGCAGCAATAGACAATTTATCTTCTATCTGGAAGAGTTTTCACA	504	
QY	1094	GATATGCAATTCGAAATATTTGTTAGTTCAGACCTATATATTTGCTGCTGAGATAG	1153	
DB	503	GATATGCAATTCGAAATATTTGTTAGTTCAGACCTATATATTTGCTGCTGAGATAG	444	
QY	1154	CTGACTCATTTGAATCAAGCCAAATTTTGGAAATTAAGTTTTCATATCTCTGTGCA	1213	
DB	443	CTGACTCATTTGAATCAAGCCAAATTTTGGAAATTAAGTTTTCATATCTCTGTGCA	384	
QY	1214	TTGTTATAGTTCCTGAGAACTGCTGGAATTTCTACTCTATTTACTTATGTCATTT	1273	
DB	383	CTGTTATAGTTCCTGAGAACTGCTGGAATTTCTACTCTATTTACTTATGTCATTT	324	
QY	1274	ATAGGCTTAAACATTAACCTGCTCCACATCCAGACTTTGTTGTAAGTGTATG	1333	
DB	323	ATAGGCTTAAACATTAACCTGCTCCACATCCAGACTTTGTTGTAAGTGTATG	264	
QY	1334	CAATGTTATTTTCATTTTTCATCTTTTCTGTAACAGACT--TTTCAGTGCCCAAT	1392	
DB	263	CAGTTGTTATTTTCATTTTTCATCTTTTCTGTAACAGACTTTTTCAGTGCCCAAT	204	
QY	1393	AGTCAGGACATCAAGGTTTATGTTGTAATATCACTGATTTTGAAGTGAAGAAATG	1452	
DB	203	AGTCAGGACATCAAGGTTTATGTTGTAATATCACTGATTTTGAAGTGAAGAAATG	144	
QY	1453	ACTTAATTA--TAGACATTTTTCACAAGAACACTGAATAGGNGGAAACATGGAATTC	1511	
DB	143	ACTTAATTAATTTAGACATTTTTCACAAGAACAGAAATAGTGGAAGAAATGGAATTC	84	
QY	1512	TTTTAGGTGACAGTGTGCTTTCATTAATTTTCTTTT--TTAATATATATTTTAACA	1570	
DB	83	TTTTAGGTGACAGTGTGCTTTCATTAATTTTCTTTTCTAATATATATTTTAACA	24	
QY	1571	TATGTAGAAATTAAGTGCAAA 1593		
DB	23	TATGTAGAAATTAAGTGCTACA 1		
RESULT 8				
AL519606	1201 bp	mRNA	linear	EST 22-MAY-2003
LOCUS	AL519606	Homo sapiens NEUROBLASTOMA COT 10-NORMALIZED	Homo sapiens	
DEFINITION	CDNA clone CS0DB004Y06 5-PRIME, mRNA sequence.			
ACCESSION	AL519606			
VERSION	AL519606.2	GI:31037953		
KEYWORDS	EST.			
SOURCE	Homo sapiens (human)			
ORGANISM	Homo sapiens			
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			

REFERENCE	Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.				
AUTHORS	1 (bases 1 to 1201)				
TITLE	L.I.W.B., Gruber,C., Jesse,J. and Polayes,D.				
JOURNAL	Full-length cDNA libraries and normalization				
COMMENT	Unpublished On Feb 13, 2001 this sequence version replaced gi:12783099. Contact: Genoscope Genoscope - Centre National de Sequencage BP 191 91006 EVRY cedex - France Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr Library was constructed by Life Technologies, a division of Invitrogen. This sequence belongs to sequence cluster 6664.f For more information about this cluster, see http://www.genoscope.cns.fr/ cgi-bin/cluster.cgi?seq=CS0DB004DE03QPL&cluster=6664.f. Contact : Feng Liang Email : fliang@life.techn.com URL : http://fulllength.invitrogen.com/ Invitrogen Corporation 1600 Faraday Avenue Genoscope sequence ID : CS0DB004DE03QPL.				
FEATURES	Location/Qualifiers				
source	1..1201 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="CS0DB004Y06" /issue_type="NEUROBLASTOMA COT 10-NORMALIZED" /note="1st strand cDNA was primed with a Molt-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."				
BASE COUNT	370 a	185 c	223 g	360 t	63 others
ORIGIN					
Query Match	31.3%;	Score 580.4;	DB 9;	Length 1201;	
Best Local Similarity	92.1%;	Pred. No. 2.5e-81;			
Matches 598;	Conservative 8;	Mismatches 42;	Indels 1;	Gaps 1;	
QY	1106	TGAATATTTGTTAGTTCAGCCTATATTTTCTGCTGGAGTATAGCTGATCTATGA	1165		
DB	66	TAAATATTTTCTTATGTTCCAGCCTATATTTTCTGCTGGAGTATAGCTGATCTATGA	125		
QY	1166	AATCAAGCCAAATTTTGGAAATTAATGTTTTCATATGCTGTCATTTGTTATATGTC	1225		
DB	126	AATCAAGCCAAATTTTGGAAATTAATGTTTTCATATGCTGTCATTTGTTATATGTC	185		
QY	1226	CTCAGAACTGCTGGAATTTGCTTACTTATTTTACCTATATGTCATTTTATGCTTATGA	1285		
DB	186	CTCAGAACTGCTGGAATTTGCTTACTTATTTTACCTATATGTCATTTTATGCTTATGA	245		
QY	1286	TAACTCTGCTCCACATCCAGACTGTTTGTGAAGTGAAGTGTGATGCAATTTGAAT	1345		
DB	246	TAACTCTGCTCCACATCCAGACTGTTTGTGAAGTGAAGTGTGATGCAATTTGAAT	305		
QY	1346	TCATAACTTTTACATCTTTTCTGAACAAGACTTTTTCAGTGCCCAATATAGTACAGACTTC	1405		
DB	306	TCATAACTTTTTCATCTTTTCTGAACAAGACTTTTTCAGTGCCCAATATAGTACAGACTTC	365		
QY	1406	AAAGGTTATGTTGTTATATGATGATATTTTGAAGTGAAGTGAAGTGAAGTGAAGTGA	1464		
DB	366	AAAGGTTATGTTGTTATATGATGATATTTTGAAGTGAAGTGAAGTGAAGTGAAGTGA	425		
QY	1465	ACCATTTCACAAAGAACAGTGAATAGGNGGAAACATGGAATTTCTTTAGTGCAGT	1524		
DB	426	ACCATTTCACAAAGAACAGTGAATAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA	485		
QY	1525	GGTGCTCTTCAATTAATTTTATATATATTTTAAATATATTTTAAATATATTAAGAATTA	1584		
DB	486	GGTGCTCTTCAATTAATTTTATATATATTTTAAATATATTTTAAATATATTAAGAATTA	545		
QY	1585	AGTGCAAGAACCTGGAAGCTTAAGACCTCTTCAANAGCTGATATATGGAAGTAA	1644		
DB	546	AGTGCAAGAACCTGGAAGCTTAAGACCTCTTCAANAGCTGATATATGGAAGTAA	605		

```

QY      1645 AANMNGTTNCAGATATCTCATATCGCTGCTNNKNNKATGAGCCCTYMNCAAGCTTGGA 1704
        ||| : ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      606 AATATGTTTCAATCATCTCATATCATCTCATATCATCTCATATCAAAAGCTTGGA 665
QY      1705 AATKTNMNTGNATAGTTNATTAANCAGGAGGCTNNMATNCT 1753
        ||| : ||| : ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      666 ATGTTTGATGATACAGATTATTAACAGGATGCTCATATTACT 714

RESULT 9
CB483386      667 bp      mRNA      linear      EST 26-MAR-2003
DEFINITION   Jns06_G08.r jns Sus scrofa cDNA 3', mRNA sequence.
ACCESSION    CB483386
VERSION      CB483386.1 GI:29289772
KEYWORDS     EST.
SOURCE       Sus scrofa (pig)
ORGANISM     Sus scrofa
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Suidae; Sus.
REFERENCE    1 (bases 1 to 667)
AUTHORS      Neilan,J.G., Kutish,G.F., Lu,Z., Zsak,A. and Rock,D.L.
TITLE        Sequence analysis of African swine fever virus infected and
              non-infected porcine macrophage cDNA libraries
              Unpublished
JOURNAL
COMMENT      Contact: Neilan JG
              Plum Island Animal Disease Center
              US Department of Agriculture, Agricultural Research Service
              PO Box 848, Greenport, NY 11944-848, USA
              Tel: 631 323 3133
              Fax: 631 323 3044
              Email: jneilane@adrc.ars.usda.gov
              Single pass sequencing. Bases called with phred v0.020425.c and
              trimmed with the aid of the trimalt option. Vector identified by
              cross_match v0.990329 and lucy v1.17p.
              Seq primer: M13 reverse.
              Location/Qualifiers
FEATURES
    source
        1..667
            /organism="Sus scrofa"
            /mol_type="mRNA"
            /db_xref="taxon:9823"
            /tissue_type="lymphoid"
            /cell_type="macrophage"
            /lab_host="DH10B"
            /note="Vector: pSPORT1; Site.1: NotI; Site.2: SalI;
            Library made from pools of polyA selected RNA. Macrophages
            were derived from peripheral blood mononuclear cells
            cultured for 48 hrs on plastic in the presence of 30% L929
            supernatant."
BASE COUNT   138 a      167 c      147 g      215 t
ORIGIN
Query Match      28.4%; Score 527.6; DB 14; Length 667;
Best Local Similarity 93.4%; Pred. No. 5.2e-73;
Matches 551; Conservative 0; Mismatches 39; Indels 0; Gaps 0;

QY      1 ATGGCGAGCTAGAGGCTTACTGTTTCTGCGCGCCTTGAGCTGACTTTTATGATCC 60
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      78 ATGGCGAGCTGAGGCTTACTTCTGCGCTGCTGAGCTGATCTTTATGATCC 137
QY      61 TGGCTCTCTTCTCCGCTTCAGCCGCGGCGCTGAGAGCCATACATGAGAGATTTTC 120
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      138 TGGCTCTCTTCTCCGCTTCAGCCGCGGCGCTGAGAGCCATACATGAGAGATTTTC 197
QY      121 CACCTGCTCAGCGCGAGCGCTACTGTGAGGCCATTTCCTCTCCAGTGGGATCC 180
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      198 CACCTGCGGAGCGCGAGCGCTACTGTGAGGCCATTTCCTCTCCAGTGGGATCC 257
QY      181 ATGATTACTACATTACTGCTGTGATCCGTGTCAGTGTGAGAGTCAACCGCCATT 240
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      258 ATGATTACTACATTACTGCTGTGATCCGTGTCAGTGTGAGAGTCAACCGCCATT 317

```

```

QY      241 TGAATCTTGCATGAGTCTGAACATGTTGTCTGCTCCATTTGGAGATGCTACATTTGTAAT 300
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      318 TGAATCTTGCATGAGTCTGAACATGTTGTGTCTGCTCCATTTGGAGATGCTACATTTGTAAT 377
QY      301 CTTCCTCTGAGTGTGGCACTTCTATTACTATATTTGCTTTTCCAGAGTACAAACC 360
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      378 CTTCCTCTGAGTGTGGCACTTCTATTATTATTTATTTGCTTTTCCAGAGTACAAACC 437
QY      361 AGAACAAGGCTGCTCAAGATATCCAGAGAGTCTGTCAACATTAACACTAGCAGATATT 420
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      438 AGAACAAGGCTGCTCAAGATATCCAGAGAGATTTTGTCAACATTAACACTAGCAGATATT 497
QY      421 CCAACACTTATTTTAACTTCCTTTATTTATACAGACAGACATCTATGTTTACT 480
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      498 CCAACATTTATTTTAACTTCCTTTATTTATACAGAAACAGGAGATCCAGTTTACT 557
QY      481 CTTTTCATATTTGATGCTGCTTTATGAAATCATTAACCTCAGCTTCCTGATTT 540
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      558 CTTTTCATATTTGATGCTGCTTTATGAAATCATTAACCTCAGCTTCCTGATTT 617
QY      541 TGTGCTTCATGTTTGGCAACAATATATCATCTGGGCTGCTCTGTCG 590
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      618 TGTGATTCATGTTTCTGCTCAACAATATCATTTGGGCTATCTTCTGTCG 667

RESULT 10
BG613309      667 bp      mRNA      linear      EST 18-APR-2001
LOCUS        602641182F1 NIH_MGC_61 Homo sapiens cDNA clone IMAGE:4771940 5',
DEFINITION   mRNA sequence.
ACCESSION    BG613309
VERSION      BG613309.1 GI:13664680
KEYWORDS     EST.
SOURCE       Homo sapiens (human)
ORGANISM     Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE    1 (bases 1 to 667)
AUTHORS      NIH-MGC http://mgi.nci.nih.gov/.
TITLE        National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL
COMMENT      Unpublished
              Contact: Robert Strausberg, Ph.D.
              Email: cgapds-remail.nih.gov
              Tissue Procurement: ATCC
              CDNA Library Preparation: CLONTECH Laboratories, Inc.
              CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
              DNA Sequencing by: Incyte Genomics, Inc.
              Clone distribution: MGC clone distribution information can be
              found through the I.M.A.G.E. Consortium/LNL at:
              http://image.llnl.gov
              Plate: ILCM1641 row: d column: 21
              High quality sequence stop: 667.
              Location/Qualifiers
FEATURES
    source
        1..667
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /clone="IMAGE:4771940"
            /tissue_type="embryonal carcinoma"
            /lab_host="DH10B (TI phage-resistant)"
            /clone_1db="NIH_MGC_61"
            /note="Organ: testis; Vector: pDNR-LIB (Clontech); Site.1:
            SfiI (ggcgctcgcc); Site.2: SfiI (ggccatcagcc);
            Double-stranded cDNA was prepared from cell line RNA. 5'
            and 3' adaptors were used in cloning as follows: 5'
            adaptor sequence: 5'-CACGGCCATTTATGCGC-3' and 3' adaptor
            sequence: 5'-ATTTCAGAGCCGAGGCGCCGACATG-dr(30)BN-3'
            (where B = A, C, or G and N = A, C, G, or T). Average
            insert size 1.75 kb (range 0.9-4.0 kb). 15/15 colonies
            contained inserts by PCR. This library was enriched for
            full-length clones and was constructed by Clontech
            Laboratories (Palo Alto, CA). Note: this is a NIH_MGC
            Library."

```

BASE COUNT 131 a 169 c 148 g 219 t
ORIGIN
Query Match 27.4%; Score 509.4; DB 10; Length 667;
Best Local Similarity 97.9%; Pred. No. 3.7e-70;
Matches 516; Conservative 0; Mismatches 11; Indels 0; Gaps 0;
QY 1 ATGGCCAGCTAGAGGATTACTGTTCTCGGCGCCTTGAGCTGATCTTTTATGATCC 60
|||||
Db 141 ATGGCCAGCTAGAGGATTACTGTTCTCGGCGCCTTGAGCTGATCTTTTATGATCC 200
QY 61 TGGCTCTCTCTTCGCGCTTCAGCGGCGGCTGCGAGAGAGCCATGAGAGATCTTC 120
|||||
Db 201 TGGCTCTCTCTTCGCGCTTCAGCGGCGGCTGCGAGAGAGCCATGAGAGATCTTC 260
QY 121 CACCTGCTCAGGCGAGGCTGCTGAGAGGCGCATTTCCCTTCCCAAGTGGATCCC 180
|||||
Db 261 CACCTGCTCAGGCGAGGCTGCTGAGAGGCGCATTTCCCTTCCCAAGTGGATCCC 320
QY 181 ATGATTACTACATTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 240
|||||
Db 321 ATGATTACTACATTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 380
QY 241 TGGATCTTGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 300
|||||
Db 381 TGGATCTTGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 440
QY 301 CTTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 360
|||||
Db 441 CTTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 500
QY 361 AGAACAAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420
|||||
Db 501 AGAACAAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 560
QY 421 CCAACCTTTATTTTAACTTCTTTATTTATTTATTTATTTATTTATTTATTTATTT 480
|||||
Db 561 CCAACCTTTATTTTAACTTCTTTATTTATTTATTTATTTATTTATTTATTTATTT 620
QY 481 CTTTTCGATTTTGTGCTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCT 527
|||||
Db 621 CTTTTCGATTTTGTGCTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCT 667
RESULT 11
B1768769 694 bp mRNA linear EST 25-SEP-2001
LOCUS 603058260F1 NIH_MGC_122 Homo sapiens cDNA clone IMAGE:5207486 5',
DEFINITION
ACCESSION B1768769
VERSION B1768769
KEYWORDS B1768769.1 GI:15760347
SOURCE EST.
ORGANISM Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 694)
NIH-MGC http://mgi.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgaabs-r@mail.nih.gov
Tissue Procurement: Life Technologies, Inc.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM1520 row: h column: 15
High quality sequence stop: 694.
Location/Qualifiers
1..694
FEATURES
source

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5207486"
/lab_host="DH10B"
/clone_lib="NIH_MGC_122"
/note="Organ: pooled lung and spleen; Vector: pCMV-Sport6;
Site 1: NotI; Site 2: EcoRV (destroyed); RNA source
anonymous pool of 24 week female lung, 16 week female
spleen, and 20-22 week male spleens. Library is oligo-dT
primed and directionally cloned (EcoRV site is destroyed
upon cloning). Average insert size 1.4 kb, insert size
range 1-3 kb. Library is normalized and enriched for
full-length clones and was constructed by C. Gruber
(Invitrogen). Research Genetics tracking code 026. Note:
this is a NIH_MGC Library."
BASE COUNT 154 a 155 c 164 g 220 t 1 others
ORIGIN
Query Match 27.1%; Score 503.4; DB 12; Length 694;
Best Local Similarity 94.0%; Pred. No. 3.2e-69;
Matches 611; Conservative 0; Mismatches 27; Indels 12; Gaps 8;
QY 1 ATGGCCAGCTAGAGGATTACTGTTCTCGGCGCCTTGAGCTGATCTTTTATGATCC 60
|||||
Db 47 ATGGCCAGCTAGAGGATTACTGTTCTCGGCGCCTTGAGCTGATCTTTTATGATCC 105
QY 61 TGGCTCTCTCTTCGCGCTTCAGCGGCGGCTGCGAGAGAGCCATGAGAGATCTT 119
|||||
Db 106 TGGCTCTCTCTTCGCGCTTCAGCGGCGGCTGCGAGAGAGCCATGAGAGATCTT 165
QY 120 CCACC-TGGCTCAGGCGAGCGCTACTGAGGCGCATTTCCCTTCCAGTGGGATC 178
|||||
Db 166 CCACC-TGGCTCAGGCGAGCGCTACTGAGGCGCATTTCCCTTCCAGTGGGATC 225
QY 179 CCATGATTTACTACATTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 238
|||||
Db 226 CCATGATTTACTACATTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 285
QY 239 TTGGGATTTGGATGCTGAGACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 298
|||||
Db 286 TTGGGATTTGGATGCTGAGACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 344
QY 299 ATCTCTCTCTACGTTGGCACTTCTATTTACTATTTGCTTTCCAAAGATCAAC 358
|||||
Db 345 ATCTCTCTCTACGTTGGCACTTCTATTTACTATTTGCTTTCTGCAAGATCAAC 404
QY 359 CCAGAAACAAGGCTGCTCAAGATTCAGAGAGCTT-GTCAACATTAACTAGAGATA 417
|||||
Db 405 CCAGAAACAAGGCTGCTCAAGATTCAGAGAGCTTGTGCAACATTAACTAGAGATA 464
QY 418 TTTCGAACCTTATTTT -TAACCTCTTATATACAGAGAGATCTATGTTT 475
|||||
Db 465 TTTCGAACCTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 524
QY 476 TTACTC-----TTTTCATATTTGATGCTGCTTATGAGAAATATATAAAGCTGCTTC 531
|||||
Db 525 TGCCTCTGATATATGCGTATGATGATGCTTATGAGAAATATATAAAGCTGCTTC 584
QY 532 CTTCGATTTTGGCTTCATGTTTGGCAACAATATATCTGCGCTGCTCTGTGCA 591
|||||
Db 585 CTTCGATTTTGGCTTCATGATGATGCGCAACAATATATCTGCGCTGCTCTGTGCA 644
QY 592 GGAATGCTATTCACAA-AGTTAACTGAGGCTTGGAAACAGACTAC 640
|||||
Db 645 GGAATGCTATTCACAAAGATTAAAGGAGGCTTGGAAACAGACTAC 694
RESULT 12
A0599053/c 560 bp DNA linear GSS 08-JUN-1999
LOCUS HS_2133_B2.B11.T7C CIT Approved Human Genomic Sperm Library D Homo
DEFINITION sapiens genomic clone Plate=2133 Col=22 Row=D, genomic survey

sequence.
ACCESSION AQ59053.1 GI:5021428
VERSION AQ59053.1
KEYWORDS GSS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Eumetazoa; Eutheria; Primates; Catarrhini; Homnidae; Homo.
AUTHORS Mahairas G.G., Wallace J.C., Smith K., Swartzell S., Holzman T., Keller A., Shaker R., Furlong J., Young J., Zhao S., Adams M.D. and Hood L.
TITLE Sequence-tagged connectors: A sequence approach to mapping and scanning the human genome
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)
MEDLINE 99380589
PubMed 10449764
COMMENT Contact: Mahairas GG, Wallace JC, Hood L
High Throughput Sequencing Center
University of Washington
401 Queen Anne Avenue North, Seattle, WA 98109, USA
Tel: (206) 616-3618
Fax: (206) 616-3887
Email: jwallaceu.washington.edu
Clones may be purchased from Research Genetics (info@resgen.com).
BAC end Web Server: <http://www.husc.washington.edu>
Plate: 2133 Row: D Column: 22
Seq primer: T7
Class: BAC ends
High quality sequence stop: 560.
FEATURES
Source Location/Qualifiers
1..560
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/clone="plate=2133 Col=22 Row=D"
/sex="male"
/clone_lib="CIT Approved Human Genomic Sperm Library D"
/note="Organ: Sperm; Vector: pBeloBAC11; BAC clones in E-Coli DH10B"
BASE COUNT 195 a 103 c 85 g 171 t 6 others
ORIGIN
Query Match 25.3%; Score 469.8; DB 28; Length 560;
Best Local Similarity 98.0%; Pred. No. 6.1e-64;
Matches 493; Conservative 0; Mismatches 8; Indels 2; Gaps 2;
QY 1107 GAATATTGTTGATCCAGCTATATATTCGTGCTGGATAGCTGATGATGAA 1166
DB 559 GAAATATTTGTTAGTCCAGCATATATATNTGCTGGTGGAGTATAGCTGATGAA 500
QY 1167 ATCAAGCAATTTTGGATTTAATGTTTTCATATGCTGTTGATTTGATAGTTC 1226
DB 499 ATCAAGCAATTTTGGATTTAATGTTTTCATATGCTGTTGATTTAGTATCC 440
QY 1227 TCAGAAATGCTGGAATTTGCTTACTTCATTTACCTTATGCTATTAAGGCTTAACAT 1286
DB 439 TCAGAAATGCTGGAATTTGCTTACTTCATTTACCTTATGCTATTAAGGCTTAACAT 380
QY 1287 AACTGCTGCCCATCCAGACTGATGTTTGGAACTGATGATGCAATGTTAATTT 1346
DB 379 AACTGCTGCCCATCCAGACTGATGTTTGGAACTGATGATGCAATGTTAATTT 320
QY 1347 CATAACTTTTACATCTTTTGAACAAGACTTTCAAGTGGCCAAATAGTCAGACATTC 1406
DB 319 CATAACTTTTACATCTTTTGAACAAGACTTTCAAGTGGCCAAATAGTCAGACATTC 260
QY 1407 AAGTTTATGTGTAATATCATGATATTTTGAAGCTGTAAAAATGACTTAATAA-TAGA 1465
DB 259 AAGTTTATGTGTAATATCATGATATTTTGAAGCTGTAAAAATGACTTAATAATA 200
QY 1466 CCATTTTACAAAGAACATGATAGNGAAACATGGAATTTCTTTTAGGTGCACTG 1525
TT

DB 199 CCATTTTACAAAGAACATGATAGTGGAAACATGCAATTTCTTTTAGGTGCACTG 140
QY 1526 GTGGCTTCGAATTTACATTTACTTTTAAATATATATTTTAAACATATGTAAAGAAATTA 1585
DB 139 GTGGCTTCGAATTTACATTTACTTTTAAATATATATTTTAAACATATGTAAAGAAATTA 81
QY 1586 GTGGCAAGAACTGGGAAGCTT 1608
DB 80 GTGGCAAGAACTGGGAAGCTT 58
RESULT 13
BX401726/c 1126 bp mRNA linear EST 13-MAY-2003
LOCUS BX401726 Homo sapiens B CELLS (RAMOS CELL LINE) COT 25-NORMALIZED
DEFINITION Homo sapiens cDNA clone CS0DL007YH07 3-PRIME, mRNA sequence.
ACCESSION BX401726
VERSION BX401726.1 GI:30618581
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Eukaryota; Eutheria; Primates; Catarrhini; Homnidae; Homo.
AUTHORS Li W.B., Gruber C., Jessee J. and Polayes D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of Invitrogen. This sequence belongs to sequence cluster 6666.f,
Contact : Feng Liang Email : liang@life.com URL : <http://fulllength.invitrogen.com/> Invitrogen Corporation 1600 Faraday Avenue Genoscope sequence ID : CS0DL007YH04NP1.
FEATURES
Source Location/Qualifiers
1..1126
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DL007YH07"
/cell_type="B CELLS (RAMOS CELL LINE) COT 25-NORMALIZED"
/clone_lib="RAMOS CELL LINE"
/note="1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."
BASE COUNT 390 a 202 c 210 g 314 t 10 others
ORIGIN
Query Match 25.0%; Score 464; DB 13; Length 1126;
Best Local Similarity 96.9%; Pred. No. 4.4e-63;
Matches 473; Conservative 0; Mismatches 15; Indels 0; Gaps 0;
QY 160 TCCCTTCCAGTGGATGCCATGATTTACATTTACCTGCTGTACTGCTGCTAGT 219
DB 497 TCAATTTTAAAGTGGATGCCATGATTTACATTTACCTGCTGTACTGCTGCTAGT 438
QY 220 GGAATGTCGAACCTGCCATTTGGATCTTTGGATGCTGTGAACATGTTGCTGCATTT 279
DB 437 GGAATGTCGAACCTGCCATTTGGATCTTTGGATGCTGTGAACATGTTGCTGCATTT 378
QY 280 GGAATGTCGAACCTGCCATTTGGATCTTTGGATGCTGTGAACATGTTGCTGCATTT 339
DB 377 GGAATGTCGAACCTGCCATTTGGATCTTTGGATGCTGTGAACATGTTGCTGCATTT 318
QY 340 CTTTTCACAAGTATCAACCCAGAAGAGCTCCTCAAGTATCCAGAGAGTCTGTCA 399
DB 317 CTTTTCACAAGTATCAACCCAGAAGAGCTCCTCAAGTATCCAGAGAGTCTGTCA 258
TT

QY 400 ACATTACACTAGACAGATTTTCCACACTTTATTTTAACTTCCTTTATATATACAGAA 459
 Db 257 ACATTACACTAGACAGATTTTCCACACTTTATTTTAACTTCCTTTATATATACAGAA 198
 QY 460 GCAGGATCTATGTTTTTACTCTTTTGCATTTTGTATGTTCTTTATGGAATCATATAA 519
 Db 197 GCAGGATCTATGTTTTTACTCTTTTGCATTTTGTATGTTCTTTATGGAATCATATAA 138
 QY 520 ACCTGACCTTCCTGATTTTGGCTTCATGTTTGGCAACAATATCATCTGGCT 579
 Db 137 ACTTACACCTTCCTGATTTTGGCTTCATGTTTGGCAACAATATCATCTGGCT 78
 QY 580 GTCCTTCTGTCAGAGGATGTCATTTGCACAAAGTTAACTGAGCTTGGAAGACTGACCTA 639
 Db 77 GTCCTTCTGTCAGAGGATGTCATTTGCACAAAGTTAACTGAGCTTGGAAGACTGACCTA 18
 QY 640 CAAAAGAA 647
 Db 17 AAGAGAGAA 10
 RESULT 14
 LOCUS AO512968 530 bp DNA linear GSS 05-MAY-1999
 DEFINITION HS_5139_A2.G04_T7A RPCI-11 Human Male BAC Library Homo sapiens
 ACCESSION AO512968 genomic clone Plate=715 Col=8 Row=M, genomic survey sequence.
 VERSION AO512968.1 GI:4745259
 KEYWORDS GSS.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 QY 1 (bases 1 to 530)
 Db Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,
 Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and
 Hood,L.
 QY Sequence-tagged connectors: A sequence approach to mapping and
 Db scanning the human genome
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)
 MEDLINE 99380589
 PUBMED 10449764
 COMMENT Contact: Mahairas GG, Wallace JC, Hood L
 High Throughput Sequencing Center
 University of Washington
 401 Queen Anne Avenue North, Seattle, WA 98109, USA
 Tel: (206) 616-3618
 Fax: (206) 616-3887
 Email: jwallace@u.washington.edu
 Library availability, please contact Pieter de Jong
 (pieter@dejong.med.buflalo.edu). Clones may be purchased from
 BACPAC Resources (http://bacpac.med.buflalo.edu/ordering_bac.htm)
 or from Research Genetics (info@resgen.com). BAC end Web Server:
 http://www.htsc.washington.edu
 Plate: 715 row: M column: 8
 Seq primer: T7
 Class: BAC ends
 High quality sequence stop: 530.
 Location/Qualifiers
 1..530
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /clone="Plate=715 Col=8 Row=M"
 /sex="male"
 /clone_lib="RPCI-11 Human Male BAC Library"
 /note="Vector: pBAC3.6; Site:1: EcoRI; Site:2: EcoRI;
 Male blood DNA was isolated from one randomly chosen donor
 and partially digested with a combination of EcoRI and
 EcoRI Methylase. Size selected DNA was cloned into the
 pBAC3.6 vector at EcoRI sites"
 BASE COUNT 120 a 91 c 92 g 222 t 5 others

ORIGIN
 Query Match 24.8%; Score 461; DB 28; Length 530;
 Best Local Similarity 93.7%; Pred. No. 1.5e-62;
 Matches 479; Conservative 0; Mismatches 32; Indels 0; Gaps 0;
 QY 678 ATTTCAGAAATTCAGAAAAATTCCTGATTTCTTTGCTTAATTCATGCTTTAAAAA 737
 Db 12 ATTGAGAGCTGTGATTCATFANCTTCGCTTCNTTGGCTTATTCATGCTTTAAAAA 71
 QY 738 CTTGAGTATGCTTTCTGTTTGTGACTTGGCCCTTACATCTTCTGGATTTCTTTTGTGC 797
 Db 72 CTTGAGTATGCTTTCTGTTTGTGACTTGGCCCTTACATCTTCTGGATTTCTTTTGTGC 131
 QY 798 TTTTGTAGTAAATGCTGGAATTTGTTATGGGATCGGAGTATGATGATGAGCTGCT 857
 Db 132 TTTTGTAGTAAATGCTGGAATTTGTTATGGGATCGGAGTATGATGATGAGCTGCT 191
 QY 858 TCATTTTCCTCACTATTTCTACTTTTTCATTTACTCTCTTTTCTTCTTCATCT 917
 Db 192 TCATTTTCCTCACTATTTCTACTTTTTCATTTACTCTCTTTTCTTCTTCATCT 251
 QY 918 CTTGCTCTCAGCAAAATTAAGACTTTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 977
 Db 252 CTTGCTCTCAGCAAAATTAAGACTTTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 311
 QY 978 TTTTGTAGTAAATGCTGGAATTTGTTATGGGATCGGAGTATGATGATGAGCTGCT 1037
 Db 312 TTTTGTAGTAAATGCTGGAATTTGTTATGGGATCGGAGTATGATGATGAGCTGCT 371
 QY 1038 CTTGCTCTCAGCAAAATTAAGACTTTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 1097
 Db 372 CTTGCTCTCAGCAAAATTAAGACTTTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 1431
 QY 1098 TGCATTTCTGAATATTTTGTAGTCCAGCCATATATTTGCTGCTGAGTATAGCTGA 1157
 Db 432 TGCATTTCTGAATATTTTGTAGTCCAGCCATATATTTGCTGCTGAGTATAGCTGA 491
 QY 1158 CTCATTTGAATCAAGCCAAATTTTGTGAAT 1188
 Db 492 CTCATTTGAATCAAGCCAAATTTTGTGAAT 522
 RESULT 15
 LOCUS BY740968 724 bp mRNA linear EST 31-DEC-2002
 DEFINITION BY740968 RIKEN full-length enriched, melanocyte Mus musculus cDNA
 ACCESSION BY740968
 VERSION BY740968.1 GI:27165555
 KEYWORDS EST.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 QY 1 (bases 1 to 724)
 Db Okazaki,Y., Furuno,M., Kasukawa,T., Adachi,J., Bono,H., Kondo,S.,
 Mikaido,I., Osato,N., Saito,R., Suzuki,H., Yamanka,I., Kiyosawa,H.,
 Yaagi,K., Tomaru,Y., Hasegawa,Y., Nogami,A., Schonbach,C.,
 Gojobori,T., Baldarelli,R., Hill,D.P., Bull,C., Hume,D.A.,
 Quackenbush,J., Schriml,L.M., Kanapin,A., Matsuda,H., Batalov,S.,
 Belset,K.W., Blake,J.A., Bradt,D., Bruschi,V., Chochia,C.F., Forrest
 L.E., Cousins,S., Dalla,E., Dragani,T.A., Fletcher,C.F., Forrest
 A., Frazer,K.S., Gaasterland,T., Gariboldi,M., Gissi,C., Godzik,A.,
 Gough,J., Grimmond,S., Gustincich,S., Hirokawa,N., Jackson,I.J.,
 Jarvis,E.D., Kanai,A., Kawai,H., Kawasawa,Y., Kedzierski,R.M.,
 King,B.L., Konagaya,A., Kurochkin,I.V., Lee,Y., Lenhard,B., Lyons
 P.A., Maglott,D.R., Maltais,L., Marchionni,L., McKenzie,L., Miki
 H., Nagashima,T., Numata,K., Okido,T., Pavan,W.J., Perlea,G.,
 Pesole,G., Petrovsky,N., Pillai,R., Pontius,J.U., Qi,D.,
 Ramachandran,S., Ravasi,T., Reed,J.C., Reed,D.J., Reid,J., Ring
 B.Z., Ringwald,M., Sandelin,A., Schneider,C., Semple,C.A., Setou
 M., Shimada,K., Sultana,R., Takenaka,Y., Taylor,M.S., Teasdale

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 23, 2003, 09:16:28 ; Search time 169 Seconds
(without alignments)
4849.991 Million cell updates/sec

Title: US-10-000-151B-1
Perfect score: 1857
Sequence: 1 atgcgcagctagagaggtta.....aagtaataatgncncaaa 1857

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/prodata/2/1na/5A_COMB.seq:*
2: /cgn2_6/prodata/2/1na/5B_COMB.seq:*
3: /cgn2_6/prodata/2/1na/6A_COMB.seq:*
4: /cgn2_6/prodata/2/1na/6B_COMB.seq:*
5: /cgn2_6/prodata/2/1na/PCUTS_COMB.seq:*
6: /cgn2_6/prodata/2/1na/backfilsl.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	502	27.0	539	3	US-09-337-171-5
2	73.4	4.0	7218	1	US-08-232-463-14
3	47.4	2.6	19124	2	US-08-487-826B-13
4	45.4	2.4	5340	4	US-09-627-122-21
5	44.6	2.4	658	4	US-08-998-416-595
6	44.6	2.4	2223	1	US-08-257-073-4
7	44	2.4	1203	4	US-09-134-001C-75
8	43.4	2.3	376	2	US-08-623-906A-18
9	43.2	2.3	2555	2	US-08-693-457-3
10	43.2	2.3	2555	2	US-09-265-731-3
11	42.8	2.3	5923	3	US-09-064-922-3
12	42.4	2.3	240	1	US-08-628-417-6
13	41.8	2.3	1020	4	US-09-328-475C-43
14	41.8	2.3	1696	4	US-09-835-811-1
15	41.8	2.3	1976	4	US-09-920-759-10
16	41.6	2.3	319608	4	US-09-679-409-1
17	41.6	2.2	6755	4	US-08-931-999-4
18	41.6	2.2	152331	3	US-09-128-155-16
19	41.6	2.2	176373	3	US-09-128-155-17
20	41.4	2.2	116592	4	US-09-818-512-3
21	41.2	2.2	454	2	US-08-623-906A-6
22	41.2	2.2	1736	3	US-09-182-816-22
23	41.2	2.2	1736	3	US-09-182-816-24
24	41.2	2.2	1736	3	US-09-471-528-22
25	41.2	2.2	1736	3	US-09-471-528-24
26	41.2	2.2	1736	3	US-09-634-530-22
27	41.2	2.2	1736	3	US-09-634-530-24

C	28	41.2	2.2	2447	2	US-09-014-969-14	Sequence 14, Appl
C	29	41.2	2.2	1664976	4	US-08-916-421B-1	Sequence 1, Appl
C	30	41	2.2	3001	4	US-09-539-333D-220	Sequence 220, App
C	31	40.8	2.2	570	4	US-09-107-532A-3368	Sequence 3368, App
C	32	40.6	2.2	319608	4	US-09-539-333D-1	Sequence 1, Appl
C	33	40.4	2.2	6124	4	US-08-213-419B-3	Sequence 3, Appl
C	34	40.4	2.2	90541	4	US-09-759-359A-3	Sequence 3, Appl
C	35	40.2	2.2	16442	3	US-08-781-891-208	Sequence 208, App
C	36	40.2	2.2	16442	4	US-09-618-166-208	Sequence 208, App
C	37	40	2.2	1493	1	US-08-340-820-24	Sequence 24, Appl
C	38	40	2.2	1493	1	US-08-593-535-24	Sequence 24, Appl
C	39	40	2.2	4285	4	US-09-410-464-1	Sequence 1, Appl
C	40	39.8	2.1	1447	4	US-09-443-041A-27	Sequence 27, Appl
C	41	39.8	2.1	5852	1	US-07-867-106-2	Sequence 2, Appl
C	42	39.6	2.1	1798	4	US-09-797-906-1	Sequence 1, Appl
C	43	39.6	2.1	4467	1	US-08-565-907A-1	Sequence 1, Appl
C	44	39.6	2.1	4467	2	US-08-910-51B-1	Sequence 1, Appl
C	45	39.6	2.1	4467	2	US-08-909-425A-1	Sequence 1, Appl

ALIGNMENTS

RESULT 1
US-09-337-171-5
; Sequence 5, Application US/09337171
; Patent No. 6262249
; GENERAL INFORMATION:
; APPLICANT: Kennedy, Giulia C
; TITLE OF INVENTION: PANCREATIC CANCER GENES
; FILE REFERENCE: 200130.454
; CURRENT APPLICATION NUMBER: US/09/337,171
; CURRENT FILING DATE: 1999-06-21
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 5
; LENGTH: 539
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-337-171-5

Query Match 27.0%; Score 502; DB 3; Length 539;
Best Local Similarity 96.8%; Pred. No. 8.2e-123;
Matches 522; Conservative 0; Mismatches 16; Indels 1; Gaps 1;

QY	1051	AATAGACATTATCTTCTATGCTGGAAGAGTTTTCAGAAATATGCAATTCGAAA	1110
DB	1	AATAGACATTATCTTCTATGCTGGAAGAGTTTTCAGAAATATGCAATTCGAAA	60
QY	1111	TATTTGTTAGTCCAGCCTATATATTTGCTGGAGTATACCTGCAATTTGAATCA	1170
DB	61	TATTTGTTAGTCCAGCCTATATATTTGCTGGAGTATACCTGCAATTTGAATCA	120
QY	1171	AAGCATTATTTTGAATTAATGTTTTCATATGCTGTTCAATTTATAGTTCGAC	1230
DB	121	AATTCATATTTTGAATTAATGTTTTCATATGCTGTTCAATTTATAGTTCGAC	180
QY	1231	AAACCTGCAATTTGTTACTTCAATTTTACCTTATATGCTTAAAGCTTAACATAC	1290
DB	181	AAACCTGCAATTTGTTACTTCAATTTTACCTTATATGCTTAAAGCTTAACATAC	240
QY	1291	CTGCGCCCAACATCCAGACTGTTTGAAGTGAAGTTCATATGCTTAAATTCATA	1350
DB	241	CTGCGCCCAACATCCAGACTGTTTGAAGTGAAGTTCATATGCTTAAATTCATA	300
QY	1351	ACTTTTACATCTTCTGGAACAAGACTTTTCAAGTGAAGTGAAGTGAAGTGAAG	1410
DB	301	ACTTTTACATCTTCTGGAACAAGACTTTTCAAGTGAAGTGAAGTGAAGTGAAG	360
QY	1411	TTTATGTTGTAATATGATGATATTTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG	1469
DB	361	TTTATGTTGTAATATGATGATATTTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG	420

OY 1470 TTCTACAAGACACTGATAGNGGAAAACATGGAATTTCTTTAGTCACTGCTG 1529
Db 421 TTCTACAAGAACACGAAATGAGTGAAGAAACATGGAATTTCTTTAGTCACTGCTG 480
OY 1530 TCTCAATATACATAGTTTATATATATTTAAACATATGTAAGAAATTAAGTG 1588
Db 481 TCTCAATATACATAGTTTATATATATTTAAACATATGTAAGAAATTAAGTG 539

RESULT 2

US-08-232-463-14
; Sequence 14, Application US/08232463
; Patent No. 5670367
; GENERAL INFORMATION:
; APPLICANT: DORNER, F.
; APPLICANT: SCHEIFLINGER, F.
; APPLICANT: FALKNER, F. G.
; TITLE OF INVENTION: RECOMBINANT FOMLOX VIRUS
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22313-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/232,463
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/935,313.
; FILING DATE:
; APPLICATION NUMBER: EP 91 114 300.6
; FILING DATE: 26-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 30472/114 IMMU
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)836-9300
; TELEFAX: (703)683-4109
; TELEX: 899149
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7218 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; CLONE: pTZ9pt-fls
; US-08-232-463-14

Query Match 4.0%, Score 73.4; DB 1; Length 7218;

Best Local Similarity 5.8%; Pred. No. 2.8e-09;

Matches 20; Conservative 207; Mismatches 118; Indels 0; Gaps 0;

OY 698 TTCTCAGTTCTTTGGCTTATTCATGTCCTTAAAACTGAGTATGCTTTCTGTT 757
Db 1123 YY 1182
OY 758 TGACTTGCCCTACATCCTCTGGGATTTCTGTTTGTCTTTGAGTACTTAATGCTG 817
Db 1183 YY 1242
OY 818 GAATGTATTTGGGAGTCGAGTAGTCAATGAAGCTGCTTCAATTTCCCACTATCTC 877
Db 1243 YY 1302

OY 878 ACTTTTTCATTTACTCTCTTTTCTTTCTTTCCCTCATCTCTCTCTCAAGAAATTA 937
Db 1303 YY 1362
OY 938 AGACTTTCTTCTCTAGTTTGGAAACATGGAATTCGTTTGTGTTACCTAGCT 997
Db 1363 YY 1422

RESULT 3

US-08-487-826B-13/C
; Sequence 13, Application US/08487826B
; Patent No. 5993827
; GENERAL INFORMATION:
; APPLICANT: Sim, Kim L.
; APPLICANT: Chitnis, Chetan
; APPLICANT: Miller, Louis H.
; APPLICANT: Peterson, David S.
; APPLICANT: Su, Xin-zhan
; APPLICANT: Wellens, Thomas E.
; TITLE OF INVENTION: BINDING DOMAINS FROM PLASMODIUM VIVAX
; NUMBER OF SEQUENCES: 45
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Knobbe Martens Olson & Bear
; STREET: 620 Newport Center Drive 16th Floor
; CITY: Newport Beach
; STATE: California
; COUNTRY: US
; ZIP: 92660
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/487,826B
; FILING DATE: 10-SEP-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Israelien, Ned
; REGISTRATION NUMBER: 29,655
; REFERENCE/DOCKET NUMBER: NIH121.001CP1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 235-8550
; TELEFAX: (619) 235-0176
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 19124 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; US-08-487-826B-13

Query Match 2.6%, Score 47.4; DB 2; Length 19124;

Best Local Similarity 46.2%; Pred. No. 0.034;

Matches 226; Conservative 0; Mismatches 261; Indels 2; Gaps 2;

OY 704 AGTTTCTTTGGCTTATTCATGTCCTTAAAACTGAGTATGCTTTCTGTTGACTT 763
Db 15879 ATTTTCTTATTTATTTATTTTATTTTAAATTTTATTTATTTATTTATTTT 15820
OY 764 GGCCCTACATCCTCTGAGATTTCTGTTTGTGCTTTTGTAGTATTAAGCTGGAATG 823
Db 15819 ATTTAAATTAATTTTATTTATTTATTTATTTATTTTATTTTAAATTTTAAATTT 15761

[illegible]

```

RESULT 4
US-09-627-122-21
: Sequence 21, Application US/09627122
: Patent No. 6472521
: GENERAL INFORMATION:
: APPLICANT: Uhlmann, Eugen
: APPLICANT: Greiner, Beate
: APPLICANT: Unger, Eberhard
: APPLICANT: Gothe, Gisliand
: APPLICANT: Schwerdel, Marc
: TITLE OF INVENTION: OLIGONUCLEOTIDES FOR THE INHIBITION OF HUMAN eg5
: TITLE OF INVENTION: EXPRESSION
: FILE REFERENCE: 02481.1678
: CURRENT APPLICATION NUMBER: US/09/627,122
: CURRENT FILING DATE: 2000-07-27
: NUMBER OF SEQ ID NOS: 23
: SOFTWARE: PatentIn Ver. 2.1
: SEQ ID NO 21
: LENGTH: 5340
: TYPE: DNA
: ORGANISM: Plasmodium falciparum
US-09-627-122-21

```

Query Match	2.48;	Score 45.4;	DB 4;	length 5340;
Best Local Similarity	46.28;	Pred. No. 0.058;		
Matches 151;	Conservative 0;	Mismatches 176;	Indels 0;	Gaps 0

[illegible]

Oy	1115	TGTTAGTTCACGCCATATTTATTTGCTGGTGGAGATAGCTGACTCATTTGAAATCAAGC	1174
Db	2188	TTTTCAGATTGCTGATATATCAAAATTTTGATTAATATCTTCACCTTGCTATTTAATAAT	2247
Oy	1175	CAATTTTGGGAATTTAATGTTTTTCA	1201
Db	2248	AAATCTTTTATTAATCTTTTATTTCCA	2274

```

RESULT 5
US-08-998-416-595
: Sequence 595, Application US/08998416
: Patent No. 6239264
: GENERAL INFORMATION:
: APPLICANT: Philippesen, Peter
: APPLICANT: Pohlmann, Rainer
: APPLICANT: Steiner, Sabine
: APPLICANT: Mohr, Christine
: APPLICANT: Wendland, Jurgen
: APPLICANT: Knechtle, Philipp
: APPLICANT: Robischung, Corinne
: TITLE OF INVENTION: GENOMIC DNA SEQUENCES OF ASHBYA GOSYPPII
: TITLE OF INVENTION: AND USES THEREOF
: NUMBER OF SEQUENCES: 1152
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: No. 6239264artis Corporation
: STREET: 3054 Cornwallis Road
: CITY: Research Triangle Park
: STATE: No. 6239264th Carolina
: COUNTRY: USA
: ZIP: 27709
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patent Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/998,416
: FILING DATE: 24-DEC-1997
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: CH 0016/97
: FILING DATE: 31-DEC-1996
: ATTORNEY/AGENT INFORMATION:
: NAME: Meigs, J. Timothy
: REGISTRATION NUMBER: 38,241
: REFERENCE/DOCKET NUMBER: PF/5-30306/A/GCCL1976
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 919-541-8587
: TELEFAX: 919-541-8689
: INFORMATION FOR SEQ ID NO: 595:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 658 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: DNA (genomic)
: ORIGINAL SOURCE:
: ORGANISM: PAG1408RP
US-08-998-416-595

```

Query Match	2.4%;	Score 44.6;	DB 3;	Length 658;
Best Local Similarity	47.3%;	Pred. No. 0.031;		
Matches 198;	Conservative	0;	Mismatches 219;	Indels 2;
			Gaps	2

[illegible]

```

; GENERAL INFORMATION:
;
; APPLICANT: Stevenson Tamar

```


APPLICANT: Dvorak, Jan
APPLICANT: Halverson, Joy
TITLE OF INVENTION: Microsatellite Sequences for Canine
TITLE OF INVENTION: Genotyping
NUMBER OF SEQUENCES: 60
CORRESPONDENCE ADDRESS:
ADDRESSEE: FLEHR, HOBBACH, TEST, ALBRITTON & HERBERT
STREET: 4 Embarcadero Center, Suite 3400
CITY: San Francisco
STATE: CA
COUNTRY: US
ZIP: 94111-4187
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/623,906A
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Sherwood, Pamela J.
REGISTRATION NUMBER: 36,677
REFERENCE/DOCKET NUMBER: A-62282/BIR
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-781-1989
TELEFAX: 415-398-3249
INFORMATION FOR SEQ ID NO: 18:
SEQUENCE CHARACTERISTICS:
LENGTH: 376 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: misc.feature
LOCATION: 1..79
OTHER INFORMATION: /note= "Nucleotides 1-79 are unique
FEATURE:
NAME/KEY: misc.feature
LOCATION: 80..229
OTHER INFORMATION: /note= "Nucleotides 80-229 are
OTHER INFORMATION: repeat sequence"
FEATURE:
NAME/KEY: misc.feature
LOCATION: 230..376
OTHER INFORMATION: /note= "Nucleotides 230-376 are
OTHER INFORMATION: unique flanking sequence"
US-08-623-906A-18
Query Match 2.38; Score 43.4; DB 2; Length 376;
Best Local Similarity 52.58; Pred. No. 0.048;
Matches 95; Conservative 0; Mismatches 86; Indels 0; Gaps 0;
QY 853 TGTCTTCATTTTCCTCAACATTTCTACTTTTTCATTACTGCTTTTTCCTTTCCT 912
Db 231 TCTTTTTCCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTT 172
QY 913 CATCTCTGCTCTCTAGCAAAATTAAGACTTTTCTTTCTTAGTTGGAACATGAAT 972
Db 171 TTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 112
QY 973 CTGTTTGGTGTACCTTAGCTCTGCTGTTTGTAGTTGGAATTCACCTATGCTCAT 1032
Db 111 CTTCCTTCTTCTTCTTCTTCTTCTTCTTCTTGTATGCTTGGTGTGTTTATAGCTCT 52
QY 1033 A 1033
Db 51 A 51

RESULT 9
US-08-693-457-3
Sequence 3, Application US/08693457
Patent No. 5880330
GENERAL INFORMATION:
APPLICANT: Weigel et al., Detlef
TITLE OF INVENTION: SHOOT MERISTEM SPECIFIC PROMOTER
TITLE OF INVENTION: SEQUENCES
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson P.C.
STREET: 4225 Executive Square, Suite 1400
CITY: La Jolla
STATE: CA
COUNTRY: USA
ZIP: 92037
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/693,457
FILING DATE: 07-AUG-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Hallie, Lisa A.
REGISTRATION NUMBER: 38,347
REFERENCE/DOCKET NUMBER: 07251/012001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619/678-5070
TELEFAX: 619/678-5099
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 2555 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-693-457-3
Query Match 2.38; Score 43.2; DB 2; Length 2555;
Best Local Similarity 46.08; Pred. No. 0.15;
Matches 132; Conservative 1; Mismatches 154; Indels 0; Gaps 0;
QY 1377 TTTTCAGTGGCCAAATAGTCAGGACATTCAAAGTTTATGCTAATATCAGTCATATT 1436
Db 641 TTTTCCTTAACAAGAGTTCATTATCTTGAGCTTGGAGTTTGAAGAGTT 700
QY 1437 TGAACGTAAATAATGAGCTTAATATAGACATTTCTACAAAGAACACTGAATAGGNG 1496
Db 701 TACATTTATTAAGAAAAATATATCATATATATATATATATATAGTATATATGAATGCT 760
QY 1497 AAAACATGAAATTTCTTTAGTGCAGTGTGCTCTTCAATTAATACATTAAGTTTAT 1556
Db 761 TTCACATTAATTTGCAACAACATCAATAATAGGTAACATACATATATAGTTTGT 820
QY 1557 ATATATTTTAACATATATGAAGAAATTAATGCGCAAGAACTGGGAAGCTTAGACCTG 1616
Db 821 TTACTCTTTAAAAAAGGAGTAACCTPAAGAGGCTATTTCTCTTAATAATTTAGCAAA 880
QY 1617 CTTCGAANGCCTGAATATAGGAAATTAAMNNGTTTNCAGATATCT 1663
Db 881 ACTGATACATGACACAAAATATGCTCATTAATTCATATTAATTTGT 927
RESULT 10
US-09-265-731-3
Sequence 3, Application US/09265731
Patent No. 6239329
GENERAL INFORMATION:
APPLICANT: Weigel et al., Detlef
TITLE OF INVENTION: SHOOT MERISTEM SPECIFIC PROMOTER
TITLE OF INVENTION: SEQUENCES

ADDRESS: DEFENSE COMMAND
 STREET: OFFICE OF THE CHIEF COUNSEL (AMSCB-GC)
 CITY: ABEDDEN PROVING GROUND
 STATE: MARYLAND
 COUNTRY: USA
 ZIP: 21010-3423
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/628,417
 FILING DATE:
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: BIFONT, ULYSSES J
 REGISTRATION NUMBER: 39,908
 REFERENCE/DOCKET NUMBER: DAM 398-94
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 410-671-1158
 TELEFAX: 410-671-2534
 INFORMATION FOR SEQ ID NO: 6:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 240 bases
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: oligodeoxynucleotide
 HYPOTHEICAL: NO
 ANTI-SENSE: YES
 US-08-628-417-6

	Query Match	2.3%	Score 42.4;	DB 1:	Length 240;
	Best Local Similarity	49.5%;	Pred.	No. 0.069;	
Matches	109;	Conservative	0;	Mismatches	111; Indels 0; Gaps 0

QY	739	TTGAGTATGCTTTTCGTGGTAGCTTGACCTTGAGCCATCATCCTTCGGATTTCGTGGTGCT	798
Dd	228	TT	169
QY	799	TTTGTAAGTAAATGCTGAATGTATAATGGCGATCGAAGTAGCAATGCATGCCGCTCT	858
Dd	168	TT	109
QY	859	CATTTCCGCACACTTCCTACTTTTTTCATTACTCTTTTTTCCTTTCCATGC	918
Dd	108	TT	49
QY	919	CTGCTCCAGCAAATAATGACCTTTCCTCCATGTT	958
Dd	48	TTTTTTTTTAGTAAAATTAATTCCTAGAAGTTTATGCTCT	9


```

RESULT 13
US-09-328-475C-43
Sequence 43, Application US/09328475C
Patent No. 6476207
GENERAL INFORMATION:
APPLICANT: Zhang, Jimmy
APPLICANT: Astrel, Jon H.
APPLICANT: Carroll III, Eddie
APPLICANT: Endege, Wilson O.
APPLICANT: Ford, Donna M.
APPLICANT: Monahan, John E.
APPLICANT: Schlegel, Robert
APPLICANT: Steinmann, Kathleen E.
TITLE OF INVENTION: GENES AND GENE EXPRESSION PRODUCTS THAT
FILE REFERENCE: ARE DIFFERENTIALLY REGULATED IN PROSTATE CANCER
CURRENT FILING DATE: 1999-06-09
NUMBER OF SEQ ID NOS: 341

```

```

: SOFTWARE: Fastseq for Windows Version 3.0.0
: SEQ ID NO 43
: LENGTH: 1020
: TYPE: DNA
: ORGANISM: Homo Sapien
: FEATURE:
: NAME/KEY: misc_feature
: LOCATION: (1)..(1020)
: OTHER INFORMATION: n = A,T,C or G
US-09-328-475C-43

```

Query Match	2.3%	Score 41.8;	DB 4;	Length 1020;
Best Local Similarity	47.7%;	Pred. No. 0.21;		
Matches 103;	Conservative 0;	Mismatches 113;	Indels 0;	Gaps 0

QY	805	GAGTACTATGGTGAATGTTATTTGGCATGGAGATAGTCATCAAGACCGTTCATATTTT	864
Db	186	GTGTGTACACAGTCATCTGTAGTACTAATATGAAGAAAGATATACAAATTTGC	245
QY	865	CCGTCAACTATTTCTACTTTTTCATTTTCACTCTCTTTTCTTTTCTCATCTCCGTCCT	924
Db	246	CCCTCAGTAGAGTACTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT	305
QY	925	CCCTAGCAAAATTAAGACTTTCTTCTTCCCTTAGTTTGGAAACATGCAATTCGTGTTTG	984
Db	306	TT	365
QY	985	GTTACCTTAGTCCTGCTGTTTGTAGTTGGAAATTC	1020
Db	366	TTTTTNGNNNTTTTNCNTTTTTTAAANAAAAAANC	401

RESULT 14

```

US-09-835-811-1/c
; Sequence 1, Application US/09835811
; Patent No. 6482936
; GENERAL INFORMATION:
; APPLICANT: HU, Song et al
; TITLE OF INVENTION: ISOLATED HUMAN SECRETED PROTEINS.
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN SECRETED PROTEINS, AND
; TITLE OF INVENTION: USES THEREOF
; FILE REFERENCE: CLO012228
; CURRENT APPLICATION NUMBER: US/09/835,811
; CURRENT FILING DATE: 2001-04-17
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 1696
; TYPE: DNA
; ORGANISM: Human
; US-09-835-811-1

```

Query Match	2.38;	Score 41.8;	DB 4;	Length 1696;
Best Local Similarity	51.38;	Pred. No. 0.28;		
Matches 97;	Conservative 0;	Mismatches 92;	Indels 0;	Gaps 0;

OY	853	IGTCGTCATTTTCCAGCACTATGCACATTTCATTCAGTCGTCTTTTCTTTCTTGCT	912
Db	1675	TT	1610
OY	913	CATCTCGCTGCTCCAGCAAATAAGACTTTCTCCTTAGTTGGAAACATGGAAAT	972
Db	1615	TT	1558
OY	973	CTGTTTTTGGTGTTACCTTAGCTCTGTGTTTTAGTTGGAAATCACATTANGCTCAT	1033
Db	1555	TTTAATCATATACAGTGATCCT	1496
OY	1033	AAATACATTG	1041
Db	1495	TAAATCTTGG	1487

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd

OM nucleic - nucleic search, using sw model

Run on: September 23, 2003, 09:16:28 ; Search time 1530 Seconds
(without alignments)
2985.665 Million cell updates/sec

Title: US-10-000-151B-1
 Perfect score: 1857
 Sequence: 1 atgycgcagctagaggtta.....aagtaatatgncncaaa 1857

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

```
Searched:      1660708 seqs, 1229959015 residues
Total number of hits satisfying chosen parameters:  3321416
```

```
Minimum DB seq length: 0
Maximum DB seq length: 20000000000
```

```
Post-processing:  Minimum Match 0%
                  Maximum Match 100%
                  Listing first 45 summaries
```

```

1:  Published.Applications.NA.*
2:  /cgn2.6/pdata/a1/pubnba/us07_PUBCOMB.seq.*
3:  /cgn2.6/pdata/a1/pubnba/PCU_NEW_PUB.seq.*
4:  /cgn2.6/pdata/a1/pubnba/us06_NEW_PUB.seq.*
5:  /cgn2.6/pdata/a1/pubnba/us06_PUBCOMB.seq.*
6:  /cgn2.6/pdata/a1/pubnba/us07_NEW_PUB.seq.*
7:  /cgn2.6/pdata/a1/pubnba/PCU5_PUBCOMB.seq.*
8:  /cgn2.6/pdata/a1/pubnba/us08_NEW_PUB.seq.*
9:  /cgn2.6/pdata/a1/pubnba/us09_PUBCOMB.seq.*
10: /cgn2.6/pdata/a1/pubnba/us09B_PUBCOMB.seq.*
11: /cgn2.6/pdata/a1/pubnba/us09C_PUBCOMB.seq.*
12: /cgn2.6/pdata/a1/pubnba/us09_NEW_PUB.seq.*
13: /cgn2.6/pdata/a1/pubnba/us10A_PUBCOMB.seq.*
14: /cgn2.6/pdata/a1/pubnba/us10B_PUBCOMB.seq.*
15: /cgn2.6/pdata/a1/pubnba/us10_NEW_PUB.seq.*
16: /cgn2.6/pdata/a1/pubnba/us06_NEW_PUB.seq.*
17: /cgn2.6/pdata/a1/pubnba/us06_PUBCOMB.seq.*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1810.6	97.5	1857	14	US-10-000-11B-1	Sequence 1, Appli
2	502	27.0	539	9	US-09-773-459-5	Sequence 5, Appli
3	485.8	16.2	910	12	US-09-814-353-21669	Sequence 21669, A
4	258.4	23.9	283	13	US-10-040-739-376	Sequence 376, A
5	71	3.8	6668	12	US-10-311-455-1670	Sequence 1670, A
6	68	3.7	14006	12	US-10-311-455-1931	Sequence 1931, A
7	65.8	3.5	16033	12	US-10-311-455-1377	Sequence 1377, A
8	64.8	3.5	9155	12	US-10-311-455-435	Sequence 435, A
9	64.2	3.5	9339	12	US-10-240-453-54	Sequence 54, Appli
10	64.2	3.5	9539	14	US-10-239-676-52	Sequence 52, Appli
11	62.6	3.4	4654	12	US-10-311-455-2196	Sequence 2196, A
12	62.2	3.3	7571	12	US-10-311-455-500	Sequence 500, Appli
13	62	3.3	17934	12	US-10-311-455-1692	Sequence 1692, A
14	61	3.3	6593	12	US-10-311-455-451	Sequence 451, Appli
15	61	3.3	6668	12	US-10-311-455-1669	Sequence 1669, A
16	60.8	3.3	6944	14	US-10-172-086-112	Sequence 112, Appli

17	60.6	3	3	18281	12	US-10-311-455-935	Sequence 935, App
18	60.6	3	3	16167	12	US-10-311-455-1056	Sequence 1056, App
19	60.6	3	3	16167	12	US-10-240-485-82	Sequence 82, App1
20	60.4	3	3	6145	12	US-10-311-455-945	Sequence 945, App
21	59.8	3	2	6334	12	US-10-311-455-1186	Sequence 1186, App
22	59.4	3	2	525	14	US-10-198-846-1483	Sequence 1483, App
23	59.2	3	2	6485	12	US-10-311-455-1781	Sequence 1781, App
24	59	3	2	5750	12	US-10-311-455-1981	Sequence 1981, App
25	58.8	3	2	6070	12	US-10-311-455-1652	Sequence 1652, App
26	58.8	3	2	6070	12	US-10-240-485-132	Sequence 132, App
27	58.8	3	2	9964	12	US-10-311-455-71	Sequence 71, App1
28	58.6	3	2	6301	12	US-10-311-455-26	Sequence 26, App1
29	58.4	3	1	12007	12	US-10-311-455-690	Sequence 690, App
30	58.2	3	1	6485	12	US-10-311-455-1782	Sequence 1782, App
31	57.6	3	1	5236	12	US-10-311-455-923	Sequence 323, App
32	57.6	3	1	8413	12	US-10-240-485-49	Sequence 49, App
33	57.4	3	1	11996	12	US-10-240-485-45	Sequence 45, App1
34	57.2	3	1	5833	12	US-10-240-453-256	Sequence 256, App
35	56.8	3	1	7657	12	US-10-311-455-1995	Sequence 1995, App
36	56.8	3	1	7657	14	US-10-239-676-185	Sequence 185, App
37	56.6	3	0	529	10	US-09-983-965-2109	Sequence 2109, App
38	56.6	3	0	5198	12	US-10-311-455-252	Sequence 252, App
39	56.4	3	0	446	10	US-09-960-352-1400	Sequence 3400, App
40	56.2	3	0	7445	12	US-10-311-455-824	Sequence 824, App
41	56.2	3	0	9539	12	US-10-240-453-33	Sequence 53, App1
42	56.2	3	0	9539	14	US-10-239-676-51	Sequence 51, App1
43	56.2	3	0	113515	12	US-10-311-455-2147	Sequence 2147, App
44	55.8	3	0	6161	12	US-10-311-455-196	Sequence 516, App
45	55.8	3	0	6960	12	US-10-311-455-2372	Sequence 2372, App

ALIGNMENTS

```

RESULT 1
US-10-000-151B-1
: Sequence 1, Application US/10000151B
: Publication No. US2003001136A1
: GENERAL INFORMATION:
: APPLICANT: Balser, Jeffrey R.
: APPLICANT: George, Alfred L.
: TITLE OF INVENTION: HUMAN KCR1 REGULATION OF HERG POTASSIUM CHANNEL BLOCK
: FILE REFERENCE: Vandebuilt Ref No. US2003001136A1 VU0120; Attorney Docket No. US22
: CURRENT APPLICATION NUMBER: US/10/000,151B
: CURRENT FILING DATE: 2000-10-30
: NUMBER OF SEQ ID NOS: 5
: SOFTWARE: PatentIn version 3.1
: SEQ ID NO 1
: LENGTH: 1857
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: misc_feature
: LOCATION: (1)..(1857)
: OTHER INFORMATION: n is any nucleic acid
: FEATURE:
: NAME/KEY: CDS
: LOCATION: (1)..(1422)
: OTHER INFORMATION: n is any nucleic acid
US-10-000-151B-1

```

Query Match	97.5%	Score 1810.6	DB 14	Length 1857
Best Local Similarity	100.0%	Pred. No. 0		
Matches 1857; Conservative	0	Mismatches	0	Indels 0; Gaps 0

Qy	Db
61	1
TGCGTCTCTTTCGCGCTTCA	ATGGCGAGCTAAGAGGGTACT
AGCCGGGCGCTGCGAGACCC	TGTTTTCGCGCGGCTTGAGCT
CTTCATGAGAGAGATTTTC	AGCTTTTTCAGCTTTTTCAGT
120	60
61	1
TGCGTCTCTTTCGCGCTTCA	ATGGCGAGCTAAGAGGGTACT
AGCCGGGCGCTGCGAGACCC	TGTTTTCGCGCGGCTTGAGCT
CTTCATGAGAGAGATTTTC	AGCTTTTTCAGCTTTTTCAGT
120	60

OY	121	NACCTGCTCAGGCGCAGCGCTACTGTGAGGGCACTTTCCTCCCTTCCAGTGGGATCC	180
Db	121	CACCTGCGCTCAGGGCGCACGGCTACTGTGAGGGCACTTTCCTCCCTTCCAGTGGGATCC	180
OY	181	ATGATTTACTACATTTCTGCGCTGTGACCTGGGTGTAGTTGGAGTGTGTAACCTGCCATT	240
Db	181	ATGATTTACTACATTTACTGCGCTTGTGACCTGGGTGTAGTTGGAGTGTGTAACCTGCCATT	240
OY	241	TGATCTTTTGGATGCTGGAACATGTTGTCTGCTCATTTGGGATGCTCAGATTTGTATAT	300
Db	241	TGATCTTTTGGATGGCTGGAACATGTTGTCTGCTCATTTGGGATGCTCAGATTTGTATAT	300
OY	301	CTTCTCTTCACTGTTTGGCACTCTCATTTACTATATTTGCTTTTCCACAAGTACACCC	360
Db	301	CTTCTCTTCACTGTTTGGCACTCTCATTTACTATATTTGCTTTTCCACAAGTACACCC	360
OY	361	AGAAAACAAGCGCTGCTCAAGTATCAGAGAGCTGTGTAACATTAACCTGTGCGATATT	420
Db	361	AGAAAACAAGCGCTGCTCAGATATCAGAGAGCTGTGTACATTTTACACTGTGCGATATT	420
OY	421	CCAACACTTTATTTTTTTTAACTCTCTTATTTATACAGAGCAGATCTATGTTTTTACT	480
Db	421	CCAACACTTTATTTTTTTTAACTCTCTTATTTATACAGAGCAGATCTATGTTTTTACT	480
OY	481	CTTTTTCGATTTTGATGTGTCTTTATGGAATCATTAACCTACGCTTCTTGATTT	540
Db	481	CTTTTTCGATTTTGATGTGTCTTTATGGAATCATTAACCTACGCTTCTTGATTT	540
OY	541	TGTGCGCTTCATGTTTGGCAAAATATCATGTCGGCTGTCTGTGCGAGGAAATGTC	600
Db	541	TGTGCGCTTCATGTTTGGCAAAATATCATGTCGGCTGTCTGTGCGAGGAAATGTC	600
OY	601	ATTGCACAAAAGTTAACTGAGGAGCTTGGAAAAGTGCATACAAAAGAAAGAGACAGACT	660
Db	601	ATTGCACAAAAGTTAACTGAGGAGCTTGGAAAAGTGCATACAAAAGAAAGAGACAGACT	660
OY	661	CCACCTATTTAAAGGACCACTTTGAGAAATTCAGAAAATTCCTTCAGTTCTTTGGCTTAT	720
Db	661	CCACCTATTTAAAGGACCACTTTGAGAAATTCAGAAAATTCCTTCAGTTCTTTGGCTTAT	720
OY	721	TCCATGTCCTTTAAAAAAGTGTAGATGTTTCTGTTGACTGTGGCCCTACATCCTGTG	780
Db	721	TCCATGTCCTTTAAAAAAGTGTAGATGTTTCTGTTGACTGTGGCCCTACATCCTGTG	780
OY	781	GGATTTCTGTTTGTCTTTTGTAGTAGTTAATGATGGAATTTGTTATGGCATGCGAGT	840
Db	781	GGATTTCTGTTTGTCTTTTGTAGTAGTTAATGATGGAATTTGTTATGGCATGCGAGT	840
OY	841	AGTATGATGAGCGCTGTCTTCACTTTTCTCTACACTATTTCACTTTTTCATTTACTCTCTT	900
Db	841	AGTATGATGAGCGCTGTCTTCACTTTTCTCTACACTATTTCACTTTTTCATTTACTCTCTT	900
OY	901	TTTTCTTTTCCCTACATCCTGTCGTCCTAGCAAAATTAAGCTTTTCTTCTCTTGTGG	960
Db	901	TTTTCTTTTCCCTACATCCTGTCGTCCTAGCAAAATTAAGCTTTTCTTCTTGTGG	960
OY	961	AAACATGGAATTCGTTTGTGGTGTACCTTAGCTCTGTGTTTATAGTTTGAATAATTC	1020
Db	961	AAACATGGAATTCGTTTGTGGTGTACCTTAGCTCTGTGTTTATAGTTTGAATAATTC	1020
OY	1021	ACTTATGCTCAATAAATCTGCTAGCAGACAAATAGACATTAATCTTCTATGTGGA	1080
Db	1021	ACTTATGCTCAATAAATCTGCTAGCAGACAAATAGACATTAATCTTCTATGTGGA	1080
OY	1081	AGAGTTTTCACAATATATGCAATTTCTGAAATTTTGTAGTTCCAGCCTATATATTGCT	1140
Db	1081	AGAGTTTTCACAATATATGCAATTTCTGAAATTTTGTAGTTCCAGCCTATATATTGCT	1140
OY	1141	GGTTGGAGTATAGCTGACTCATTTGAAATCAAAAGCAATTTTGGAAATTTAATGTTTTTC	1200
Db	1141	GGTTGGAGTATAGCTGACTCATTTGAAATCAAAAGCAATTTTGGAAATTTAATGTTTTTC	1200
OY	1201	ATATGCTGTTCAATGTTATATAGTTCCTCAGAAAGTGTGGAATTTGTTACTTTCATTTTA	1260

Db	1201	AAATCTGTTTCATTTAGTTTAACTTCCTCGAAAACAGCTCGGAATTTGGTATCTTCATTTTA	1260
QY	1261	CCCTATGTCATTTATAGGCTTAACATTAATCTGCGCCGCCACATCCAGACTGTTGTGAA	1320
Db	1261	CCTTATGTCATTTATAGGCTTAACATTAATCTGCGCCGCCACATCCAGACTGTTGTGAA	1320
QY	1321	CTGAGTTCGTATGCAATTTGTTAAATTTTCATTAACCTTTTTCATCTTCTCGAACAGACTTTT	1380
Db	1321	CTGAGTTCGTATGCAATTTGTTAAATTTTCATTAACCTTTTTCATCTTCTCGAACAGACTTTT	1380
QY	1381	CAGTGGCCAAATAGTCAGACATTCGAAGGTTTATGTGTAATATAGTATATTTTGA	1440
Db	1381	CAGTGGCCAAATAGTCAGACATTCGAAGGTTTATGTGTAATATAGTATATTTTGA	1440
QY	1441	CTGTAAATATGCACTTAATTAATAGACCAATTTCTACAAACACAACTGAATAGGAGAAA	1500
Db	1441	CTGTAAATATGCACTTAATTAATAGACCAATTTCTACAAACACAACTGAATAGGAGAAA	1500
QY	1501	CATGGAATTTCTTTTAGGTGACAGTGGTGTCTTCAATTAATTAATATAT	1560
Db	1501	CATGGAATTTCTTTTAGGTGACAGTGGTGTCTTCAATTAATTAATATAT	1560
QY	1561	ATTTTAAACATATGTAAGAAATTAAGTGGCAAGAACTGGGAAAGCTTAAGACTGCTTC	1620
Db	1561	ATTTTAAACATATGTAAGAAATTAAGTGGCAAGAACTGGGAAAGCTTAAGACTGCTTC	1620
QY	1621	AAANCCCGAATTAATGGGAAATTAANNNMTTNCAGATATCCATATGCTCCTNNKNAT	1680
Db	1621	AAANCCCGAATTAATGGGAAATTAANNNMTTNCAGATATCCATATGCTCCTNNKNAT	1680
QY	1681	GNTGGCCCTTNNCAANGCTTGGGAATGKTNTNMNTGNATAGTTNATTAANCTGGGNT	1740
Db	1681	GNTGGCCCTTNNCAANGCTTGGGAATGKTNTNMNTGNATAGTTNATTAANCTGGGNT	1740
QY	1741	GCTNNMATNACTTNNNNKNCANCCWNNNNMACNATGNNNTANNANTATTTACAAAGCTC	1800
Db	1741	GCTNNMATNACTTNNNNKNCANCCWNNNNMACNATGNNNTANNANTATTTACAAAGCTC	1800
QY	1801	AGGTATATCTTGTGACTGAGAAAGTGCCTCCTNNAACATTAAGTAAATATGAGCCCAAA	1857
Db	1801	AGGTATATCTTGTGACTGAGAAAGTGCCTCCTNNAACATTAAGTAAATATGAGCCCAAA	1857
RESULT 2			
US-09-773-459-5			
Sequence 5, Application US/09773459			
Patent No. US20010016651A1			
GENERAL INFORMATION:			
APPLICANT: Kennedy, Giulia C			
TITLE OF INVENTION: PANCREATIC CANCER GENES			
FILE REFERENCE: 200130.454			
CURRENT APPLICATION NUMBER: US/09/773,459			
CURRENT FILING DATE: 2001-01-31			
PRIOR APPLICATION NUMBER: US 09/337,171			
PRIOR FILING DATE: 1999-06-21			
NUMBER OF SEQ ID NOS: 15			
SOFTWARE: FastSeq for Windows Version 3.0			
SEQ ID NO 5			
LENGTH: 539			
TYPE: DNA			
ORGANISM: Homo sapien			
US-09-773-459-5			
Query Match			
Best Local Similarity 96.8%; Score 502; DB 9; Length 539;			
Matches 522; Conservative 0; Mismatches 16; Indels 1; Gaps 1;			
QY	1051	AATAGACATTAATCTTCTATCTGTGGAAAAAGTTTTCGAAGAATATGCATTTGGA	1110
Db	1	AATAGACATTAATCTTCTATCTGTGGAAAAAGTTTTCGAAGAATATGAACCTGTA	60
QY	1111	TATTTGTTAGTCCAGCCTATATATTTGCTGTGGAGATATGCTGACATTTGAATCA	1170

|||||
Db 61 TATTGTAGTCCAGCTATATATTGCTGGAGTATACCTGATTAATCA 120
QY 1171 AAGCCATTTTGGAAATTAATGTTTCATATGCTGTTCATTTAGTCTCAG 1230
Db 121 AAGTCAATTTTGGAAATTAATGTTTCATATGCTGTTCATTTAGTCTCAG 180
QY 1231 AAGCTGTGAATTTGTTTACTTCAATTTTACCTATGTCATTTATGAGCTTAACATAC 1290
Db 181 AAGCTGTGAATTTGTTTACTTCAATTTTACCTATGTCATTTATGAGCTTAACATAC 240
QY 1291 CTGCTCCACATCCAGACTTGTGTGAAGTGTGCTATGCTATGCTTAATTTCA 1350
Db 241 CTGCTCCACATCCAGACTTGTGTGAAGTGTGCTATGCTATGCTATGCTTAATTTCA 300
QY 1351 ACTTTTACATCTTTTGAACAGACTTTTCACTGGCCAAATGATGAGCATTCAGAG 1410
Db 301 ACTTTTACATCTTTTGAACAGACTTTTCACTGGCCAAATGATGAGCATTCAGAG 360
QY 1411 TTTATGTGTAATATCAGTATTTTGAAGTGTGAAGTGTGCTATGCTTAATTTCA 1469
Db 361 TTTATGTGTAATATCAGTATTTTGAAGTGTGAAGTGTGCTATGCTTAATTTCA 420
QY 1470 TTTTACAAAGACACTGAATGAGGAAACATGGAATTTCTTTAGTGTGAGTGTG 1529
Db 421 TTTTACAAAGACACTGAATGAGGAAACATGGAATTTCTTTAGTGTGAGTGTG 480
QY 1530 TCTTCAATTTACCTTTTATATATATTTTAAACATATGTAAGAAATTAAGTG 1588
Db 481 TCTTCAATTTACCTTTTATATATATTTTAAACATATGTAAGAAATTAAGTG 539

RESULT 3

US-09-814-353-21669/c
; Sequence 21669, Application US/09814353
; Publication No. US20030165831A1
; GENERAL INFORMATION:
; APPLICANT: Lee, John
; APPLICANT: Thompson, Pamela
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; FILE REFERENCE: MRI-006B
; CURRENT APPLICATION NUMBER: US/09/814,353
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/191,031
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: US 60/207,124
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 60/211,940
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: US 60/216,820
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: US 60/220,661
; PRIOR FILING DATE: 2000-07-25
; PRIOR APPLICATION NUMBER: US 60/257,672
; PRIOR FILING DATE: 2000-12-21
; NUMBER OF SEQ ID NOS: 22037
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 21669
; LENGTH: 910
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 909, 910
; OTHER INFORMATION: n = A,T,C or G
US-09-814-353-21669

Query Match 26.28; Score 485.8; DB 12; Length 910;
Best Local Similarity 94.38; Pred. No. 3.3e-107;
Matches 516; Conservative 0; Mismatches 27; Indels 4; Gaps 1;

QY 1 ATGGCCAGCTAGAGGATTTACTGTTCTGCGCCGCTTGAGTGTACTTTTATGTC 60
Db 787 ATGGCCAGCTAGAGGATTTACTGTTCTGCGCCGCTTGAGTGTACTTTTATGTC 728
QY 61 TGGCTCTCTCTCTGCGCCCTTACCGCGGCTGCGAGAGGCTTACATGAGCATCTTC 120
Db 727 TGGCTCTCTCTCTGCGCCCTTACCGCGGCTGCGAGAGGCTTACATGAGCATCTTC 668
QY 121 CACCTGCTCAGCGCAGCGCTTACTGTGAGGCGCATTTTCTCCCTTCCAGTGGATCC 180
Db 667 CACCTGCTCAGCGCAGCGCTTACTGTGAGGCGCATTTTCTCCCTTCCAGTGGATCC 608
QY 181 ATGATTACATTAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 240
Db 607 ATGATTACATTAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 548
QY 241 TGGATCTTGGATGCTGTGAACATGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 300
Db 547 TGGATCTTGGATGCTGTGAACATGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 488
QY 301 CTTCCTCTCAGTGTGCGCACTTCTATTTATTTGCTTTCCAGAGTACAAACC 360
Db 487 CTTCCTCTCAGTGTGCGCACTTCTATTTATTTGCTTTCCAGAGTACAAACC 428
QY 361 AGAACAAGGCTGCTCAAGTATCCAGAGAGTCTGTCAACATTAACACTAGCATATT 420
Db 427 AGAACAAGGCTGCTCAAGTATCCAGAGAGTCTGTCAACATTAACACTAGCATATT 368
QY 421 CCAACACTTATTTTAACTCTTATTTATATAGAGAGCAGATCTATGTTTACT 480
Db 367 CCAACACTTATTTTAACTCTTATTTATATAGAGAGCAGATCTATGTTTACT 312
QY 481 CTTTTGCATTTTATGATGCTCTTATGGAATCATTAACCTCAGCTTCTGATTT 540
Db 311 ACTCTTCTGATTTATGATGCTCTTATGGAATCATTAACCTCAGCTTCTGATTT 252
QY 541 TGTGCT 547
Db 251 TGTGCT 245

RESULT 4

US-10-040-739-376
; Sequence 376, Application US/10040739
; Publication No. US20020173635A1
; GENERAL INFORMATION:
; APPLICANT: Jacobs, Kenneth
; APPLICANT: McCoy, John
; APPLICANT: Lavallee, Edward
; APPLICANT: Racie, Lisa
; APPLICANT: Merberg, David
; APPLICANT: Treacy, Maurice
; APPLICANT: Spaulding, Vikki
; TITLE OF INVENTION: SECRETED, EXPRESSED SEQUENCE TAGS
; NUMBER OF SEQUENCES: 1519
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genetics Institute, Inc.
; STREET: 87 CambridgePark Drive
; CITY: Cambridge
; STATE: Massachusetts
; COUNTRY: U.S.A
; ZIP: 02140
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy Disk
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/040,739
; FILING DATE: 07-Jan-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:

```

US-10-311-455-1670

Query Match      3.8% Score 71; DB 12; Length 6668;
Best Local Similarity 44.0%; Pred. No. 2.6e-06;
Matches 302; Conservative 0; Mismatches 385; Indels 0; Gaps 0;

QY 698 TTTCTCAGTTCCTTTGGCTATTCACATGTCCTTAATAAACTGAGATAGCTTTTCGT 757
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2766 TTTTTTGTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 2825

QY 758 TGACTTGGCCCTACATCCTTCTGGATTTCGTTTGTGCTTTTGTAGTAGTAATGNG 817
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2826 TTTTTTTTGGTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 2885

QY 818 GAATGTATTTGGCGAATGAGATGATGATGAAGCGCTCTCATTTTCCCAACTATCT 877
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2886 TTTTTTTTGGTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 2945

QY 878 ACCTTTTTCATTTACCTCTTTTCTTTCTTTCTCTCACCTGCCTCCAGCAAAATTA 937
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2946 TTTTTTTTGGTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 3005

QY 938 AGACTTTTCTTCTTCCCTAGTTTGGAAACATGSAATCTGTTTTGNGTTACCCTACT 997
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3006 TTTTTTTTGGTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 3065

QY 998 CTGTGTTTTAGTTTGAATAATCATTATGCTCAATAATCTGTCTGACAGACAATAGAC 1057
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3066 TTTTTTTTGGTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 3125

QY 1058 ATTATCTTCTATGCTGSGAAAAGACTTTTTTCAAATATATGCAATCTCGAATAATTT 1117
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3126 TTTTTTTTGGTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 3185

QY 1118 TAGTCCAGCCCTATATATTTGCTGTTGAGATAGCTGACTATGAAATCAAGCCAA 1177
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3186 TTTTTTTTGGTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 3245

QY 1178 TTTTGTGAAATTAATGTTTTTCATATGCTGTTGATGTTAATAGTCTCAGAACTGC 1237
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3246 TTTTTTTGTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTG 3305

QY 1238 TCGAATTTCTGTTACTTATTTTACCTATATGCTATTTATAGGCTTAACATACCTGCC 1297
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3306 TTTTTTTTGGTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 3365

QY 1298 CCACATCCAGACTGTTTGGAACTAGTGCATGCAATGTTAATTTATTAACCTTTT 1357
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3366 TTCTGATATTTAAGGGAATGAAAGTGTCTGAATTTTAAATATTTGAATATTTGATTTT 3425

QY 1358 ACAATCTTCTGAAACAGACTTTTCAGT 1384
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3426 TAATTTTCGAAAGAATGATGTTTGAT 3452

RESULT 6
US-10-311-455-1931
? Sequence 1931, Application US/10311455
? Publication No. US20030143606A1
? GENERAL INFORMATION:
? APPLICANT: OLEK, Alexander
? APPLICANT: PIEPENBROCK, Christian
? APPLICANT: BERLIN, Kurt
? TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Det
? TITLE OF INVENTION: cytosine methylation
? FILE REFERENCE: 5013.1014
? CURRENT APPLICATION NUMBER: US/10/311.455
? CURRENT FILING DATE: 2002-12-16
? PRIOR APPLICATION NUMBER: PCT/EP01/07537
? PRIOR FILING DATE: 2001-07-02
? PRIOR APPLICATION NUMBER: DE 10032529.7
? PRIOR FILING DATE: 2000-06-30
? PRIOR APPLICATION NUMBER: DE 10043826.1

```


;; PRIOR FILING DATE: 2000-09-01
;; NUMBER OF SEQ ID NOS: 2424
;; SEQ ID NO 1931
;; LENGTH: 14006
;; TYPE: DNA
;; ORGANISM: Artificial Sequence
;; FEATURE:
;; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
;; NAME/KEY: unsure
;; LOCATION: 8289, 8310, 8313
;; OTHER INFORMATION: n is a or g or c or t
US-10-311-455-1931

Query Match 3.7%; Score 68; DB 12; Length 14006;
Best Local Similarity 44.8%; Pred. No. 2.2e-05;
Matches 260; Conservative 0; Mismatches 320; Indels 0; Gaps 0;

QY 698 TTCTCAGTTTCCTTTGGCTTATTCATGCTTTAAAACTTGAGTATGCTTTCTGTT 757
DB 1924 TTTTCTTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 1983
QY 758 TGACTTGGCCCTACATCTTCCTGGATTTCTGTTTGTGCTTTGTAGTAAAGTNG 817
DB 1984 TTTTCTTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 2043
QY 818 GAATTTGATTTGGCATCGAGTATGATGAACCGTCTTCATTTTCCGACATATCT 877
DB 2044 TTTTCTTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 2103
QY 878 ACTTTTTCATTTTACTCTCTTTTTCCTTTTCCATCTCTGCTCTAGCAAAATTA 937
DB 2104 TTTTCTTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 2163
QY 938 AGACTTTCTTCCTTACGTTTGAACAATGGAATTCGTTTGTGGTGTTACTTACTCT 997
DB 2164 TTTTCTTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 2223
QY 998 CTCTGTTTATTTAGTTTGAATTCATTTATGCTATTAATATCTGCTAGCAATATGAC 1057
DB 2224 TTTTCTTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 2283
QY 1058 ATTATACCTTCTATGTTGGAAAAAGATTTTCAAGATATGCAATTCGAAATATTTGT 1117
DB 2284 TTTTCTTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 2343
QY 1118 TAGTTCACGCTATATTTTCTGTTGGATATGCTGACATTCGAAATCAAGCCAA 1177
DB 2344 TTTTCTTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 2403
QY 1178 TTTTGTGAATTTAATGTTTTCATATGCTTGTTCATTTGTATAGTTCGCAAACTGC 1237
DB 2404 TTTTCTTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 2463
QY 1238 TGGAAATTTCTTACTTCAATTTTACCTTATGCTATTTATAG 1277
DB 2464 TTTTCTTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 2503

RESULT 7
US-10-311-455-1377
; Sequence 1377, Application US/10311455
; Publication No. US20030143606A1
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Detect
; FILE REFERENCE: 5013.1014
; CURRENT APPLICATION NUMBER: US/10/311,455
; CURRENT FILING DATE: 2002-12-16
; PRIOR APPLICATION NUMBER: PCT/EP01/07537

;; PRIOR FILING DATE: 2001-07-02
;; PRIOR APPLICATION NUMBER: DE 10032529.7
;; PRIOR FILING DATE: 2000-06-30
;; PRIOR APPLICATION NUMBER: DE 10043826.1
;; PRIOR FILING DATE: 2000-09-01
;; NUMBER OF SEQ ID NOS: 2424
;; SEQ ID NO 1377
;; LENGTH: 16033
;; TYPE: DNA
;; ORGANISM: Artificial Sequence
;; FEATURE:
;; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-311-455-1377

Query Match 3.5%; Score 65.8; DB 12; Length 16033;
Best Local Similarity 47.3%; Pred. No. 8.2e-05;
Matches 199; Conservative 0; Mismatches 222; Indels 0; Gaps 0;

QY 698 TTCTCAGTTTCCTTTGGCTTATTCATGCTTTAAAACTTGAGTATGCTTTCTGTT 757
DB 3071 TTTTCTTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 3130
QY 758 TGACTTGGCCCTACATCTTCCTGGATTTCTGTTTGTGCTTTGTAGTAAAGTNG 817
DB 3131 TTTTCTTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 3190
QY 818 GAATTTGATTTGGCATCGAGTATGATGAACCGTCTTCATTTTCCGACATATCT 877
DB 3191 TATTTTCTTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 3250
QY 878 ACTTTTTCATTTTACTCTCTTTTTCCTTTTCCATCTCTGCTCTAGCAAAATTA 937
DB 3251 TTTTCTTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 3310
QY 938 AGACTTTCTTCCTTACGTTTGAACAATGGAATTCGTTTGTGGTGTTACTTACTCT 997
DB 3311 TTTTCTTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 3370
QY 998 CTCTGTTTATTTAGTTTGAATTCATTTATGCTATTAATATCTGCTAGCAATATGAC 1057
DB 3371 TTTTCTTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 3430
QY 1058 ATTATACCTTCTATGTTGGAAAAAGATTTTCAAGATATGCAATTCGAAATATTTGT 1117
DB 3431 ATTAGGAGAGAGGTGTTGTAAGAGGATTTAGAAATGCTTATATAGTATTTAT 3490
QY 1118 T 1118
DB 3491 T 3491

RESULT 8
US-10-311-455-435
; Sequence 435, Application US/10311455
; Publication No. US20030143606A1
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Det
; FILE REFERENCE: 5013.1014
; CURRENT APPLICATION NUMBER: US/10/311,455
; CURRENT FILING DATE: 2002-12-16
; PRIOR APPLICATION NUMBER: PCT/EP01/07537
; PRIOR FILING DATE: 2001-07-02
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 2424
; SEQ ID NO 435
; LENGTH: 9155

```

: TYPE: DNA
: ORGANISM: Artificial Sequence
: FEATURE:
: OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-311-455-435

Query Match
Best Local Similarity 50.2%; Score 64.8; DB 12; Length 9155;
Pred. No. 0.0001;
Matches 212; Conservative 0; Mismatches 207; Indels 3; Gaps 2;

QY      853  TGCCTCATTTTCCCTCAACTATCTACTTTTTTTCATCTTACTCTCTTTTCCCT 912
Db       1916  TGAATTTATTTTATTTATTTTATTTGTTTTTTTTTATTTATTTATTTTATTTT
QY      913  CATCTCCGTCTCCTCAAAATTAAGACTTTTCTTCCCTAGTTTGGAAACATGAAAT 972
Db       1976  TTTTATTTTATTAATATAGAAATATATAGGTTTTTATATAGGGAATTTAGTTGGG 2035
QY      973  CTGTTTTTGGTGTTACCTTAGTCGTGTGTTTTTAACTTTGGAAATTCACATGCTAT 1032
Db       2036  TTACTTTTATTTAGTTTTATTTAGTTT-TATTTTATTTTATTTAGAAATAATGTTTATTT
QY      1033  AAATACCTTGCAGAGACAATAGCATATATACCTTTATGCTGTGAGAAAGAGTTTCCA 1092
Db       2095  GGAATTTATTTAAATATTTTATTAAGATTTAATTCATTTATTTAAATTTTTTTTAAAT 2154
QY      1093  AGATATGCAATTCGAAATATTTGTTA--GTTCAGCCCTATATTTTCTGCTGTGAGTA 1150
Db       2155  AAATATTTAATTTATTTATTTTATTTTAAAGTATATGAAAAAATGTTGTAATAGAAATTT 2214
QY      1151  TAGCTGACTCATTTAAATCAAGCCAAATTTTTTGGAAATTTAATGTTTTCATATGCTTGT 1210
Db       2215  GATTTTATTTTAAAGAAAGAAAGATTTTGTGGTTTTTTTTTATTTTATTTTGTGT 2274
QY      1211  TCATTTGTTATTTAGTCTCTCAGAAACTGCTGCAATTTCTTACTTATTTTACCTATGTCA 1270
Db       2275  TTAGCTTTAGGCTTTTATTTAATGTTGAACGTTGTTTGTTTAAATTTTGTATATTTA 2334
QY      1271  TT 1272
Db       2335  GT 2336

RESULT 9
US-10-240-453--54
: Sequence 54, Application US/10240453
: Publication No. US20030148326A1
: GENERAL INFORMATION:
: APPLICANT: OLEK, Alexander
: APPLICANT: PIEPENBROCK, Christian
: APPLICANT: BERLIN, Kurt
: TITLE OF INVENTION: Diagnosis of Diseases Associated with DNA
: TITLE OF INVENTION: Transcription
: TITLE OF INVENTION: by Means of Assessing the Methylation Status of Genes Associated
: FILE REFERENCE: 5013.1009
: CURRENT APPLICATION NUMBER: US/10/240.453
: PRIOR APPLICATION NUMBER: PCT/EP01/03973
: PRIOR FILING DATE: 2001-04-06
: PRIOR APPLICATION NUMBER: DE 10019058.8
: PRIOR FILING DATE: 2000-04-06
: PRIOR APPLICATION NUMBER: DE 10019173.8
: PRIOR FILING DATE: 2000-04-07
: PRIOR APPLICATION NUMBER: DE 10032529.7
: PRIOR FILING DATE: 2000-06-30
: PRIOR APPLICATION NUMBER: DE 10043826.1
: PRIOR FILING DATE: 2000-09-01
: NUMBER OF SEQ ID NOS: 350
: SEQ ID NO 54
: LENGTH: 9539
: TYPE: DNA
: ORGANISM: Artificial Sequence

```

```

: FEATURE INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-240-453-54

Query Match      3.5%, Score 64.2; DB 12; Length 9539;
Best Local Similarity 45.7%; Pred. No. 0.00015;
Matches 263; Conservative 0; Mismatches 308; Indels 4; Gaps 1

OY 698 TTTCTCAGTTCTTTGGCTATTCCATGCTCTTAATAAAGCTGAGTAATGCTTTCTGT 757
Db 360 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 419
OY 758 TGACTTGGCCCTACATCCCTCTGGATTTCTGTTTGTCTTTGTACTAGTAATGCG 817
Db 420 TTTTCTTGTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTG 479
OY 818 GAATTTGTTATGGGATCGAGTACGTATGAAAGCCTGCTTCATTTTCCATCAATTCY 877
Db 480 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTG 539
OY 878 ACTTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 937
Db 540 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 599
OY 938 AGACTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 997
Db 600 ATATATTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 659
OY 998 CTGCTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1055
Db 660 TGTGTTAGTGTGAGCTGTACTGCTACATTTCTGTTTATTTATTTTCTTTTCTGT 719
OY 1054 AGACATTATATCTTCTATGTGTGGAAGAGCTTTTCTCAAGATATGCAATTCGAAT 1113
Db 720 AGTATTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCT 779
OY 1114 TTTGTTAGTTCAGGCTATATTTGTCGTGTGAGTATAGCTATAGCTATGGAATCAAG 1172
Db 780 TGGTTAGTTTGTGTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 839
OY 1174 CCAATTTCTTGGAAATTAATGTTTCTTCAATGCTTTGTCATTTAGTTCCTCAGAA 1233
Db 840 TTGAATTTTCTGTTTAAAGTCATTTTCTTATTTAGTTTCTTAAAGTATGGAATTAT 899
OY 1234 CTGCTGCAATTCGTACTTCTTTTACCTTATGCT 1268
Db 900 GTGTGAGTATATTTTCGTTTTTTTTTTTTTGT 934

RESULT 10
US-10-239-676-52
: Sequence 52, Application US/10233676
: Publication No. US20030082609A1
: GENERAL INFORMATION:
: APPLICANT: OLEK, Alexander
: APPLICANT: PIEPENBROCK, Christian
: APPLICANT: BERLIN, Kurt
: TITLE OF INVENTION: Diagnosis of Diseases Associated with Gene Regulation
: FILE REFERENCE: 5013.1003
: CURRENT APPLICATION NUMBER: US/10/239.676
: CURRENT FILING DATE: 2002-09-24
: PRIOR APPLICATION NUMBER: PCT/EP01/03968
: DE 10019058.8
: DE 10019173.8
: DE 10032529.7
: DE 10043826.1
: PRIOR FILING DATE: 2001-04-06
: 2000-04-06
: 2000-04-07
: 2000-06-30
: 2000-09-01
: NUMBER OF SEQ ID NOS: 228
: SEQ ID NO 52

```

LENGTH: 9539
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-239-676-52

Query Match 3.5%; Score 64.2; DB 14; Length 9539;
Best Local Similarity 45.7%; Pred. No. 0.00015;
Matches 263; Conservative 0; Mismatches 308; Indels 4; Gaps 1;

QY 698 TTCTTCAGTTCTTTGGCTTAATTCAGTCCTTTAAACCTGAGTATGCTTTCTGT 757
DB 360 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 419
QY 758 TGAATGGCCCTACACCTCTGAGATCTGCTTTGCTTTGCTTTGCTTTGCTTTGCT 817
DB 420 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 479
QY 818 GAATGTATATGCGATCGAGTATGATGACCTGCTTCATTTTCTCAACTATTTCT 877
DB 480 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 539
QY 878 ACTTTTTCATTTACTCTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCT 937
DB 540 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 599
QY 938 AGACTTTCTTCTCTAGTATGGAACATGGAATCTGCTTTGCTTTGCTTTGCTTTGCT 997
DB 600 ATTTATTTTATTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 659
QY 998 CTGTGT---TTTATGTTTGAATTCACCTTATGCTCAATTAATCTTGTCTAGACAT 1053
DB 660 TGTGTAGGTGAGAGTATGAGTATGAGTATGAGTATGAGTATGAGTATGAGTAT 719
QY 1054 AGCATTTATCTTCTCTATGCTGAGAAAGATTTTCAAGATATGCAATTCGAATAT 1113
DB 720 AGGATTTTCTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 779
QY 1114 TTTTCTTTCTCAGCTATATATTTGCGGTGAGTATGCTTATGCTTATGCTTATGCT 1173
DB 780 TGTGTAGTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 839
QY 1174 CCAATTTTGAATTTAATGTTTTCATATGCTTGTCTTATGTTATGTTATGCTTCA 1233
DB 840 TTTAATTTTGTGTTTAAAGTATTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 899
QY 1234 CTGCTGAATTTGCTTACTTCAATTTTCACTTATGT 1268
DB 900 GTGTGAGTTATTTATTTTCTGCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 934

RESULT 11
US-10-311-455-2196
Sequence 2196, Application US/10311455
Publication No. US20030143606A1
GENERAL INFORMATION:
APPLICANT: OLEK, Alexander
APPLICANT: PIEPENBROCK, Christian
APPLICANT: BERLIN, Kurt
TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Detecting
FILE REFERENCE: 5013.1014
CURRENT APPLICATION NUMBER: US/10/311.455
CURRENT FILING DATE: 2002-12-16
PRIOR APPLICATION NUMBER: PCT/EP01/07537
PRIOR FILING DATE: 2001-07-02
PRIOR APPLICATION NUMBER: DE 10032529.7
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: DE 10043826.1
PRIOR FILING DATE: 2000-09-01
NUMBER OF SEQ ID NOS: 2424
SEQ ID NO 2196

LENGTH: 4654
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-311-455-2196

Query Match 3.4%; Score 62.6; DB 12; Length 4654;
Best Local Similarity 42.9%; Pred. No. 0.00023;
Matches 379; Conservative 0; Mismatches 495; Indels 10; Gaps 1;

QY 684 AGAATTCAGAAATATTCAGTATTTCTTTGGCTTAATTCAGTCCTTTAAACCTGAG 743
DB 1517 ATATTTTATTAATAATTTATTTATTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 1576
QY 744 TATGCTTTTCTTTGCTTTGCTTTGCTTTGCTTTGCTTTGCTTTGCTTTGCTTTGCT 803
DB 1577 GGAATTTTCTTTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTAT 1636
QY 804 AGATTTAATGTCGATTTGTTATTTGCGATGAGTATGATGACCTGCTTTCATTT 863
DB 1637 TGTATTTATTCGATTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 1696
QY 864 TCCGCAATATTCATTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 923
DB 1697 TTTTGGGCTTTTAAATTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 1756
QY 924 TCCGCAATATTTCACTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 983
DB 1757 GAATGTGCTTTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 1816
QY 984 GGTATCTTATGCTCTGCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCT 1043
DB 1817 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 1876
QY 1044 AGCAGCAATATGACATTTATTTCTTATGCTGGAAGAAAGATTTTCAAGATATGCAAT 1103
DB 1877 TATATATTTATGAGTATGAGTATGAGTATTTTAAATGATATTTTCTTTTCTTTTCTTT 1936
QY 1104 TCGAATATTTGTTAGTATCCAGCTTATATTTCTGCTGCTGAGTATGCTGATTT 1163
DB 1937 TATGATTTTATTTGTTGTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 1986
QY 1164 GAATCAAAAGCAATTTTGAATTTATTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 1223
DB 1987 TTTTATTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 2046
QY 1224 TCCGCAAAAGCTGCAATTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 1283
DB 2047 TATGTTGTAATAATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 2106
QY 1284 CATACCTGCTCCCTCCCAATCCAGACTTTTGTGAACGTAGTGTGATGCAATTTGTTAA 1343
DB 2107 TATGGAATTTATTTAGAGAGTTTATTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 2166
QY 1344 TTTTCAATTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 1403
DB 2167 TTTAGGTTTATTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 2226
QY 1404 TCAAGGTTTATTTGTAATATTCAGTATTTTGAACGTGTAATAATGACTTAATATA 1463
DB 2227 TTTATATATGTTTATTTTATTTTGAATATTTTATGATGTAATATTTGTTATTTTAA 2286
QY 1464 GACATTTCTACAAAGCAACGATATGAGCAAAACATGCAATTTCTTTTGTGCTCAG 1523
DB 2287 AGTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 2346
QY 1524 TGTGTGCTTCAATTTCAATTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 1567
DB 2347 TTTATTTTCTTAATAAGTTTATTTGGAATATTTAGTATTTAGTATTTAGTATTTAGT 2390

RESULT 12

Db 16065 TTTTGTGATTTTAAATTTTTTTTTTTTATTTTAAATTTTTTATTTTAAAT 16124
QY 1373 AGACTTTTCAGTCGCCAAATAGTCAGACATTCACAAAGTTTATGCTGATATACAGAT 1432
Db 16125 ATGTTTATTTTATTTAGTAGATAGAGATTTTAAATTTATTTAGTGTTTGTGATTTT 16184
QY 1433 ATTTGAACTGTAATAATGAGACTTAATTAATAGACATTTCTACAAAGACACTGAATG 1492
Db 16185 GTCGTTTTTAAATTTTGGGTTTTTAAATGCTGTTTTTATTTTATTTTATTTT 16244
QY 1493 GNGAAACATGGAATTTCTTTAGTGCGAGTGCTGCTCAATTTACATTTAGTTT 1552
Db 16245 GGAATTTTATTTTATTTTAAATTTTATTTTATTTTATTTTGAAGTTTTTTTGA 16304
QY 1553 TAATATATTTTAAACATATCTAA 1577
Db 16305 TAATTTATTTTATTAATAGATTAA 16329

RESULT 14

US-10-311-455-451
; Sequence 451, Application US/10311455
; Publication No. US20030143606A1
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Determining Cytosine Methylation
; FILE REFERENCE: 5013.1014
; CURRENT APPLICATION NUMBER: US/10/311,455
; CURRENT FILING DATE: 2002-12-16
; PRIOR APPLICATION NUMBER: PCT/EP01/07537
; PRIOR FILING DATE: 2001-07-02
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 2424
; SEQ ID NO 451
; LENGTH: 6593
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-311-455-451

Query Match 3.3%; Score 61; DB 12; Length 6593;

Best Local Similarity 42.8%; Pred. No. 0.00069;
Matches 364; Conservative 0; Mismatches 481; Indels 5; Gaps 1;

QY 744 TATGCTTTCTGTTGACCTGGCCCTACATCTTCGGGATTTCTGTTGCTTTGCTTTGCT 803
Db 164 TTTTGTGTTTTTTTTTTTGTGTTGTTTTTTTTTGTGTTGTTGATTTTGTGATG 223
QY 804 AGTAGTTAATGCGAATGTTATTTGGCATCGAGAGTACATGAAGCCGCTGCTATT 863
Db 224 GATATGTTTGTATTTTATTTTATTTTGTATGATGATATTTTATTTTATTTT 283
QY 864 TCCCTCACTAATCT 923
Db 284 TTTATGTAATGATTTTATTTTATGTTTATGTTTATGTTTATTTTATTTTATTTT 343
QY 924 TCCCTACAAATTAACATTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 978
Db 344 GATATTTATTAATTAATTAATTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 403
QY 979 TTGCTGTTACCTTACCTGCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 1038
Db 404 TTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 463
QY 1039 TTGCTGACGACAAATAGACATTAATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1098

Db 464 GTTTGAGCTGATTAATTTTTTTTGTGATTAATTTGTTGTTAAATTTTTTATTTATGATGA 523
QY 1099 GCAATTCGTAATATTTGTTAGTTCACGCTATATATTTGCTGTTGAGATATAGCTGAC 1158
Db 524 TTTTATTTTATTTATTTGATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 583
QY 1159 TCAATGAATTCAAAGCAATTTTGTGAATTTAATGTTTTCATATGCTGTTCAATGTT 1218
Db 584 TTTTATTTTATTTTATTAATTTTATTTTATTTAATGCTATTTTATTTATTTGCTTGA 643
QY 1219 ATGTTCTCAGAACTGCTGCAATTTGCTTACTTATTTTACCTATATGCTATTTAGG 1278
Db 644 GTTGTATTTATTTGTTTGTGTTTGTAAATTTGTTGATTTATTAAGATGATTTATTTGA 703
QY 1279 CTTAACATACCTGCTGCTCCACATCGACACTGTTTGTGAAGCTGATGCTATTCGAAT 1338
Db 704 TTTTGTGATTAATTTATATGATTTTATTTTATTTTATTTTATTTTATTTTATTTT 763
QY 1339 GTTAATTTCAATCTTTTATACATCTTTCTGACACACTTTTCAGTGCCAAATAGTCAG 1398
Db 764 GTTTTATTTGATTCGTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 823
QY 1399 GACATTCAAAGTTTATGCTGTAATATTCAGTATTTTGAAGCTGTAATAAGTACTTAA 1458
Db 824 GTTTTATATGATATTTATGTTTAAATGATATGTTTATTTTATTTTATTTTATTTATG 883
QY 1459 TAATACACATTTCTCAAAAGAACACTGAATGAGGAAACATGGAATTTCTTTAGG 1518
Db 884 TGATAGGAAAGATGTTATGATGATGTTTGGTTTGTGTTTGTGTTTATTTATTTATTTAG 943
QY 1519 TGCAGTGTGCTGCTCAATTTACATTTGTTTATATATATTTTAAACATATGTAAG 1578
Db 944 TGTATGTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 1003
QY 1579 AAATTAAGTC 1588
Db 1004 TAATTTTGTG 1013

RESULT 15

US-10-311-455-1669
; Sequence 1669, Application US/10311455
; Publication No. US20030143606A1
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Determining Cytosine Methylation
; FILE REFERENCE: 5013.1014
; CURRENT APPLICATION NUMBER: US/10/311,455
; CURRENT FILING DATE: 2002-12-16
; PRIOR APPLICATION NUMBER: PCT/EP01/07537
; PRIOR FILING DATE: 2001-07-02
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 2424
; SEQ ID NO 1669
; LENGTH: 6668
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
; NAME/KEY: unsure
; LOCATION: 4733
; OTHER INFORMATION: n is a or g or c or t
US-10-311-455-1669

Query Match 3.3%; Score 61; DB 12; Length 6668;

Best Local Similarity 49.8%; Pred. No. 0.00069;

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 23, 2003, 09:16:28 ; Search time 22785 Seconds
(Without alignments)
3334.175 Million cell updates/sec

Title: US-10-000-151b-1
Perfect score: 1857
Sequence: 1 atggcgacgactagaggtta.....aagtaatatngcncnaa 1857

Scoring table: IDENTITY_NNC
Gapop 10.0 , Gapext 1.0

Searched: 2888711 segs, 20454813386 residues
Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*

1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_ov:*
5: gb_ov:*
6: gb_ov:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vl:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rod:*
36: em_htg_man:*
37: em_htg_vtl:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1647.4	88.7	2390	9 HSA312278	AJ312278 Homo sapi
2	1645.4	88.6	2280	6 AX498580	AX498580 Sequence
3	1388.2	85.5	1820	9 BC033730	BC033730 Homo sapi
4	1359.8	73.2	1461	6 AX591660	AX591660 Sequence
5	1349.2	72.7	165301	9 AC117372	AC117372 Homo sapi
6	1349.2	72.7	179032	2 AC024000	AC024000 Homo sapi
7	1311.8	70.6	1489	6 AX591658	AX591658 Sequence
8	1298.4	69.9	176600	9 AC046130	AC046130 Homo sapi
9	1233.8	66.4	158647	2 AC024478	AC024478 Homo sapi
10	1217.2	65.5	2362	6 BD160531	BD160531 Primer fo
11	1217.2	65.5	2362	9 AK027657	AK027657 Homo sapi
12	1075	57.9	7628	10 RNU78090	U78090 Rattus norv
13	775.6	41.8	276372	2 AC125569	AC125569 Rattus no
14	746.8	40.2	152534	2 AC114560	AC114560 Mus muscu
15	725.2	39.1	807	6 BD148636	BD148636 Primer fo
16	534	28.8	208006	2 AC087877	AC087877 Mus muscu
17	343.2	18.5	469	6 AX079565	AX079565 Sequence
18	336.4	18.1	442	6 AX079646	AX079646 Sequence
19	258.4	13.9	283	6 BD060016	BD060016 Secreted
20	202.4	10.9	2043	3 AK112724	AK112724 Ciona int
21	174.6	9.4	856	6 BD148633	BD148633 Primer fo
22	174.6	9.4	2245	6 BD158186	BD158186 Primer fo
23	174.6	9.4	2245	9 AK023061	AK023061 Homo sapi
24	151.8	8.2	546	6 BD153420	BD153420 Primer fo
25	88.6	4.8	1541	3 AY128497	AY128497 Drosophi
26	86.8	4.7	120142	2 AC017890	AC017890 Drosophi
27	86.8	4.7	161175	3 AC010023	AC010023 Drosophi
28	86.8	4.7	281592	3 AE003546	AE003546 Drosophi
29	86.4	4.7	1524	3 DME431376	AJ431376 Drosophi
30	83.6	4.5	41327	3 CBR647J19	AC084676 Caenorhab
31	80.8	4.4	93791	2 AC138073	AC138073 Homo sapi
32	77.6	4.2	66993	2 AC138074	AC138074 Homo sapi
33	76	4.1	111882	2 AC115612	AC115612 Dictyoste
34	74.6	4.0	1576	8 BT002117	BT002117 Arabidops
35	74.6	4.0	1758	8 AY099540	AY099540 Arabidops
36	73.4	4.0	7218	6 I66494	I66494 Sequence 14
37	72.6	3.9	8750	3 AY028643	AY028643 Plasmodiu
38	72	3.9	6644	6 E23356	E23356 Virus vecto
39	72	3.9	7372	6 E23357	E23357 Virus vecto
40	72	3.9	7797	6 E23355	E23355 Virus vecto
41	72	3.9	7996	6 E23359	E23359 Virus vecto
42	71	3.8	6668	6 AX346599	AX346599 Sequence
43	71	3.8	289973	2 AC135678	AC135678 Rattus no
44	70.4	3.8	143585	2 AC013349	AC013349 Homo sapi
45	70.4	3.8	156550	2 AC015830	AC015830 Homo sapi

ALIGNMENTS

RESULT 1
LOCUS HSA312278
DEFINITION Homo sapiens mRNA for alpha2-glucosyltransferase (ALG10 gene).
ACCESSION AJ312278
VERSION AJ312278.1 GI:14349124
KEYWORDS ALG10 gene; alpha2-glucosyltransferase.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1
AUTHORS Burda, P. and Aebi, M.
TITLE The ALG10 locus of Saccharomyces cerevisiae encodes the alpha-1,2
glucosyltransferase of the endoplasmic reticulum.

glucose of the lipid-linked oligosaccharide is required for efficient N-linked glycosylation
JOURNAL Glycobiology 8 (5), 455-462 (1998)
MEDLINE 98259831
PUBMED 9597543
AUTHORS 2
TITLE Oriol, R., Moore, S., Chantret, I., Mollicon, R. and Codogno, P.
REFERENCE Common origin and evolution of glycosyltransferases using
DOI-P-monosaccharides as donor substrates
JOURNAL Unpublished
AUTHORS Oriol, R.
TITLE 3 (bases 1 to 2390)
REFERENCE Direct Submission
AUTHORS Submitted (07-JUN-2001) Oriol R., US04, INSERM, 16 Av. Paul
JOURNAL Vaillant-Couturier, Villejuif, 94807 CEDEX, FRANCE
FEATURES
source
1..2390
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="12"
/tissue_type="whole embryo"
/dev_stage="embryo"
/country="France"
1..2390
/gene="ALG10"
320..1741
/gene="ALG10"
/function="synthesis of N-glycans"
/codon_start=1
/evidence="experimental"
/product="alpha2-glycosyltransferase"
/protein_id="CAC41349.1"
/db_xref="GI:14349125"
/translation="MAOEGYFSAALSCFLVSLCLFSAFSAIRLREPMDEIHLPO
AORCEGFHESIQMDPMITTLPGLYIVSIGIKRAIMFEGSEHVCISGIMREINPL
FVGNFYLLYLLECKVQPRKASISIQVLSLTLAVPTLFFENLVTGASKEFT
LFAIMCLYGNHKTSAFLGFCGEMRQINIIWAVYCAENVIAOKLTEMKTELORKE
RLPIKGFPAERKILQFLAYSMFKMLSMILLTWPTLIGFLFCFVVVNGIIV
GDRSHSEACHLPOLIFYEFSEFLFESPHLSKIKFLSLVMKRLIFVYVLVSV
FLVMKTEYAHKYLLADNRHYEYVWAKRVFORERYKYLLVAVIYFAGSIDSLSKRS
IFAMLMFICLFETIVIPDKLLEPRYFILPIYIRLNIPPLPSRLICLSYAVVNF
TEFTLNKTPQPMPSQDIQRFMW"
repeat_region
214..2181
/rpt_family="Alu repeat"
/rpt_unit="214..2181"
polyA_signal
2359..2364
/gene="ALG10"
BASE COUNT 635 a 476 c 456 g 823 t
ORIGIN

Query Match 88.7%; Score 1647.4; DB 9; Length 2390;
Best Local Similarity 95.8%; Pred. No. 0;
Matches 1680; Conservative 7; Mismatches 66; Indels 1; Gaps 1;

QY 1 ATGGGCACTAGAGGCTTACTGTTCTCGGCCGCTTACGCTGACTTTTACTGTC 60
DB 320 ATGGGCGAGCTGAGAGGCTTACTGTTCTCGGCCGCTTACGCTGACTTTTACTGTC 379
QY 61 TGCCCTCTCTCTCGGCTTACGCGGGGCTGCGAGAGCCCTACATGAGAGATCTTC 120
DB 380 TGCCCTCTCTCTCGGCTTACGCGGGGCTGCGAGAGCCCTACATGAGAGATCTTC 439
QY 121 CACCTGCTCAGGCGCAGGCTTACTGTTGAGGGCCATTTCCCTTCCAGTGGATGCC 180
DB 440 CACCTGCTCAGGCGCAGGCTTACTGTTGAGGGCCATTTCCCTTCCAGTGGATGCC 499
QY 181 ATGATTAATCAATTAATGCTTGTACCTGTGTACAGTGTGATGAGTGAACCTGCACAT 240
DB 500 ATGATTAATCAATTAATGCTTGTACCTGTGTACAGTGTGATGAGTGAACCTGCACAT 559
QY 241 TGGATCTTTGGATGCTGAGACATGTTGTCTGCTCATTTGGATGCTGAGATTTGTAAAT 300
DB 1640 TGGATCTTTGGATGCTGAGACATGTTGTCTGCTCATTTGGATGCTGAGATTTGTAAAT 1699

DB 560 TGGATCTTTGGATGCTGAGACATGTTGTCTGCTCATTTGGATGCTGAGATTTGTAAAT 619
QY 301 CTTCCTCAGAGTGGGCAACTTCTATTACTATATATTTGCTTCCAGAGGTACACCC 360
DB 620 CTTCCTCAGAGTGGGCAACTTCTATTACTATATTTGCTTCCAGAGGTACACCC 679
QY 361 AGAAACAAGGCTGCTCAAGATCCAGAGAGCTTGTCAACATTAACACTAGACATATT 420
DB 680 AGAAACAAGGCTGCTCAAGATCCAGAGAGCTTGTCAACATTAACACTAGACATATT 739
QY 421 CCAACACTTATTTTAACTCTCTTATATATACAGAGAGATATGTTTTTACT 480
DB 740 CCAACACTTATTTTAACTCTCTTATATATACAGAGAGATATGTTTTTACT 799
QY 481 CTTTGGCATTTTATGATGCTTATATGAAATCATAAACTTCAGCTTCTTGATTT 540
DB 800 CTTTGGCATTTTATGATGCTTATATGAAATCATAAACTTCAGCTTCTTGATTT 859
QY 541 TGTGCTTCATGTTTGGCAACAATATCATCTGGCTGTCTGTCAGAGAAATGTC 600
DB 860 TGTGCTTCATGTTTGGCAACAATATCATCTGGCTGTCTGTCAGAGAAATGTC 919
QY 601 ATTGCACAAAGTTACTGAGGCTTGAAGAACTGAGTCACAAAGAGAGACAGACTT 660
DB 920 ATTGCACAAAGTTACTGAGGCTTGAAGAACTGAGTCACAAAGAGAGACAGACTT 979
QY 661 CCACCTATTAAAGACCAATTGACAAATTCAGAAATTTCTTCACTTTTGGCTTAT 720
DB 980 CCACCTATTAAAGACCAATTGACAAATTCAGAAATTTCTTCACTTTTGGCTTAT 1039
QY 721 TCCATGTCCTTTAAACCTTGATATGCTTTTCTGTTGACTTGCCCTACATCTCTG 780
DB 1040 TCCATGTCCTTTAAACCTTGATATGCTTTTCTGTTGACTTGCCCTACATCTCTG 1099
QY 781 GGATTTCTGTTTGGCTTTGCTTGTAGTAAATGTTGGAATTTGAGTGGAGT 840
DB 1100 GGATTTCTGTTTGGCTTTGCTTGTAGTAAATGTTGGAATTTGAGTGGAGT 1159
QY 841 AGTCATGAAGCTGCTTCAATTTCTCAACATATCTTTTCTTCACTTACTCTCTT 900
DB 1160 AGTCATGAAGCTGCTTCAATTTCTCAACATATCTTTTCTTCACTTACTCTCTT 1219
QY 901 TTTTCTTTCTCATCTCTCTGCTCTGCTGACAAATTAAGCTTTTCTTCTTCTTCTT 960
DB 1220 TTTTCTTTCTCATCTCTCTGCTCTGCTGACAAATTAAGCTTTTCTTCTTCTTCTT 1279
QY 961 AAACATGAATTCGTTTGTGTTGCTTACCTGCTGCTGCTGTTTGTGTTGGAATTC 1020
DB 1280 AAACATGAATTCGTTTGTGTTGCTTACCTGCTGCTGCTGTTTGTGTTGGAATTC 1339
QY 1021 ACTATGCTCAATAATTAATCTGCTAGCAGACATATACATTAATCTTCTATGTTGAAA 1080
DB 1340 ACTATGCTCAATAATTAATCTGCTAGCAGACATATACATTAATCTTCTATGTTGAAA 1399
QY 1081 AGAGTTTTCAAAGATATGCAATTCGAAATTTTGTAGTTCCAGCCTATATATTGCT 1140
DB 1400 AGAGTTTTCAAAGATATGCAATTCGAAATTTTGTAGTTCCAGCCTATATATTGCT 1459
QY 1141 GGTGGAGATATAGGCTCAATTCGAAATTCGAAATTTTGTGAAATTTAATGTTTTC 1200
DB 1460 GGTGGAGATATAGGCTCAATTCGAAATTCGAAATTTTGTGAAATTTAATGTTTTC 1519
QY 1201 ATATGCTTGTTCATTTATATAGTTCTTCAGAAAGCTGAGAAATTTCTTACTTCAATTTA 1260
DB 1520 ATATGCTTGTTCATTTATATAGTTCTTCAGAAAGCTGAGAAATTTCTTACTTCAATTTA 1579
QY 1261 CCTATATGCTATTTATAGCTTAAACATTAATCTGCTCCACATCCAGACTGTTTGTGAA 1320
DB 1580 CCTATATGCTATTTATAGCTTAAACATTAATCTGCTCCACATCCAGACTGTTTGTGAA 1639
QY 1321 CTGAGTGTCTATGCAATTTGTAATTTTCACTTTTTCATCTTTCTGAAACAGACTTT 1380
DB 1640 CTGAGTGTCTATGCAATTTGTAATTTTCACTTTTTCATCTTTCTGAAACAGACTTT 1699

QY 1261 CCTATGTCATTATAGGCTTAAACATACCTGCTCCACATCCAGACTTGTGTGAA 1320
|||||
Db 1354 CCTATGTCATTATAGGCTTAAACATACCTGCTCCACATCCAGACTTGTGTGAA 1413
QY 1321 CTGAGTGCATVAGCAATGTTAAATTCATACCTTTTACATCTTTGACACAGACTTT 1380
|||||
Db 1414 CTGAGTGCATVAGCAATGTTAAATTCATACCTTTTACATCTTTTACACAGACTTT 1473
QY 1381 CAGTGGCCAAATAGTCAGGACATTCAAAGTTTATGCTGATATACAGTATATTGAA 1440
|||||
Db 1474 CAGTGGCCAAATAGTCAGGACATTCAAAGTTTATGCTGATATACAGTATATTGAA 1533
QY 1441 CTGTAATAATGACTTAATAA-TAGACCATTTCTCAAGAACAACATGATAGNGGAAA 1499
|||||
Db 1534 CTGTAATAATGACTTAATAA-TAGACCATTTCTCAAGAACAACATGATAGNGGAAA 1593
QY 1500 ACATGGAATTTCTTTTATGCTGAGTGTGCTTCAAAATACATTAGTTTTTAAATA 1559
|||||
Db 1594 ACATGGAATTTCTTTTATGCTGAGTGTGCTTCAAAATACATTAGTTTTTAAATA 1653
QY 1560 TATTTTAACATATCTAGAAATTAAGCGCAAGAAGTGGAAAGCTTAGACCTGCTT 1619
|||||
Db 1654 TATTTTAACATATCTAGAAATTAAGCGCAAGAAGTGGAAAGCTTAGACCTGCTT 1713
QY 1620 CAANAGCCTGATATGCGGAAATTAANMNGTTTNCAGATATCTATATGCTCCTNNKNA 1679
|||||
Db 1714 CAANAGCCTGATATGCGGAAATTAANMNGTTTNCAGATATCTATATGCTCCTCATA 1773
QY 1680 TGTGTCGCCCTTMCANAGCTTGGAGATGTTNNMNTGATATGTTTAAATCTGGGNN 1739
|||||
Db 1774 TGTGTCGCCCTTMCANAGCTTGGAGATGTTNNMNTGATATGTTTAAATCTGGGTA 1833
QY 1740 TGTCTNNMATNA 1751
|||||
Db 1834 TGTCTCAAAAAA 1845

RESULT 3
BC033730 1820 bp mRNA 1linear PRI 08-JUL-2002
LOCUS Homo sapiens, clone MGC:4485 IMAGE:5168319, mRNA, complete cds.
DEFINITION BC033730
ACCESSION BC033730.1 GI:21707441
VERSION MGC.
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 1820)
AUTHORS Strausberg, R.
TITLE Direct Submission
JOURNAL Submitted (02-JUL-2002) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: <http://mgc.ncl.nih.gov>
Contact: MGC help desk
Email: cgapbs-rt@mail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: National Institutes of Health Intramural
Sequencing Center (NISC),
Gaithersburg, Maryland;
Web site: <http://www.nisc.nih.gov/>
Contact: nisc-mgc@ncl.nih.gov
Ahtler, N., Ayale, K., Beckstrom-Sternberg, S.M., Benjamin, B.,
Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S.,
Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P.,
Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Latic, P., Legaspi, R.,
Machuro, Q.L., Mastello, C., Maskerl, B., Mastrian, S.D., McCloskey, J.C.,
McDowell, J., Pearson, R., Stanlipop, S., Thomas, P.J., Touchman, J.W.,

Tsurgeon, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L.,
Young, A., Zhang, L.-H. and Green, E.D.
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LNL at: <http://image.lnl.gov>
Series: IRAC Plate: 68 Row: 1 Column: 12
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 14249543.
FEATURES
location/Qualifiers
source
1. 1820
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="MGC:4485 IMAGE:5168319"
/tissue_type="Brain, adult medulla"
/clone_id="NIH_MGC_119"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6"
101. 1522
/codon_start=1
/product="unknown (protein for MGC:4485)"
/protein_id="AAH33730.1"
/db_xref="gi:21707442"
CDS
1.
/translation="MAOLEGYESAALSTFVSCILFSAFSRALRPIYDEIRHLP
AQRCEGHFSLSDPMTTLPGLYLSIGIRPAIMFGSEHVCSIGMRVNL
FSVNGFYLILYLFCKVOPRNKAASSIORVSLTFLAVFPLTFEFLYEAQSMFET
LEAYLMCIYGNHKTSAFLGFCGFEMROTNILIAVFCAGNIMIAQTEAMKELOKED
RLPPIKGFPAERKILQFLAYSMSPKLSMLILTPYILGFLFCAFYVNGIYV
GDRSHACHPOLFLYFESFTLPSFKILSPKILFSLVWKRRILFVPLVAV
FLVKKFTYAHKYLADNRHITYFYWKRYFOYEVKYLIVPAVYIFAMSTADSKS
IFNMFLPKTCLFVIVPOKLEFRFIFLPIVYIRLNIPLPRTSLICELSYAVVNI
TFEFLNPFQPMNSODIQRPM"

BASE COUNT 504 a 348 c 334 g 634 t
ORIGIN

Query Match 85.5%; Score 1588.2; DB 9; Length 1820;
Best Local Similarity 97.2%; Pred. No. 0;
Matches 1622; Conservative 1; Mismatches 44; Indels 1; Gaps 1;

QY 1 ATGGGCGAGTAAAGAGGTTACTGTTTCTCGGCGCCCTTGAGCTGACTTTTATGTC 60
|||||
Db 101 ATGGGCGAGTAAAGAGGTTACTGTTTCTCGGCGCCCTTGAGCTGACTTTTATGTC 160
QY 61 TGCTCTCTCTCTCCGCTTACGCGGGGCGTGGAGAGCCCTACAGGACAGATCTTC 120
|||||
Db 161 TGCTCTCTCTCTCCGCTTACGCGGGGCGTGGAGAGCCCTACAGGACAGATCTTC 220
QY 121 CACCTGCTCAGGCGGACGCTACTGTGAGGCGCATTTCTCCCTTCCAGTGGAGATCC 180
|||||
Db 221 CACCTGCTCAGGCGGACGCTACTGTGAGGCGCATTTCTCCCTTCCAGTGGAGATCC 280
QY 181 ATGATTACTACATTACCGGCTTGTACTGTGTCAGTGGAGGTCGTAACCTGCCATT 240
|||||
Db 281 ATGATTACTACATTACCGGCTTGTACTGTGTCAGTGGAGGTCGTAACCTGCCATT 340
QY 241 TGGATCTTGGATGGTGTGAACATGTGTCTGCTCATTGGAGTGTGATTTGTTAT 300
|||||
Db 341 TGGATCTTGGATGGTGTGAACATGTGTCTGCTCATTGGAGTGTGATTTGTTAT 400
QY 301 CTTCCTTCATGTTGGCACTTCTATTACTATATTTGTTTCCACAAGGTACACCC 360
|||||
Db 401 CTTCCTTCATGTTGGCACTTCTATTACTATATTTGTTTCCACAAGGTACACCC 460
QY 361 AGAACAAGGCTGCTCAAGTATCCAGAGAGTCTGTCAACATTAACTAGAGATATT 420
|||||
Db 461 AGAACAAGGCTGCTCAAGTATCCAGAGAGTCTGTCAACATTAACTAGAGATATT 520
QY 421 CCAACACTTATTTTAACTCTCTTATATACAGAGAGATCTATGTTTAACT 480
|||||
Db 521 CCAACACTTATTTTAACTCTCTTATATACAGAGAGATCTATGTTTAACT 580
QY 481 CTTTTCATATTTGATGTGCTTTATATGGAATCATAAACTTACAGCTTCTTGATTT 540
|||||

Db 581 CTTTGGCGTATTGATGTCCTTTATGGAATCATAAACCTTCACGCTTCTTGGATTT 640
QY 541 TGTGCTTCATGTTTGGCAACAATATCATCTGGCTGTCTTGTGCGAGGAATGTC 600
Db 641 TGTGCTTCATGTTTGGCAACAATATCATCTGGCTGTCTTGTGCGAGGAATGTC 700
QY 601 ATTGCAAAAAGTTAACTGAGCGCTTGGAAAACTGACCTACAAAAGAAGACAGACTT 660
Db 701 ATTGCAAAAAGTTAACTGAGCGCTTGGAAAACTGACCTACAAAAGAAGACAGACTT 760
QY 661 CCACCTATTAAAGACCATTTTGCAGAAATTCAGAAAAATCTTGTAGTTCTTTGGCTTAT 720
Db 761 CCACCTATTAAAGACCATTTTGCAGAAATTCAGAAAAATCTTGTAGTTCTTTGGCTTAT 820
QY 721 TCCATGCTCCTTTAAACCTTGAATGATGCTTCTGTTTGTGCTGGCCCTACATCTCTTG 780
Db 821 TCCATGCTCCTTTAAACCTTGAATGATGCTTCTGTTTGTGCTGGCCCTACATCTCTTG 880
QY 781 GGATTTCTGTTTGTGCTTGTAGTATGTTAATGTTGAATTTTATGGCGATCGAGT 840
Db 881 GGATTTCTGTTTGTGCTTGTAGTATGTTAATGTTGAATTTTATGGCGATCGAGT 940
QY 841 AGCATGAAGCCCTGCTTCTATTTCTCTCAACTATCTACTTTTTCATTTACTCTCTT 900
Db 941 AGCATGAAGCCCTGCTTCTATTTCTCTCAACTATCTACTTTTTCATTTACTCTCTT 1000
QY 901 TTTTCTTTTCTCATCTCCCTGCTCCTAGCAAAATTAAGACTTTCTCTTCTAGTTTG 960
Db 1001 TTTTCTTTTCTCATCTCCCTGCTCCTAGCAAAATTAAGACTTTCTCTTCTAGTTTG 1060
QY 961 AAACATGAATTCGTGTTTGTGTTGATGTTTACCTGATGCTGTTTGTAGTTTGAATTC 1020
Db 1061 AAACATGAATTCGTGTTTGTGTTGATGTTTACCTGATGCTGTTTGTAGTTTGAATTC 1120
QY 1021 ACTATGCTCATTAATACCTGCTGAGACAGACATATGATATCTTCTATGTTGTGAAA 1080
Db 1121 ACTATGCTCATTAATACCTGCTGAGACAGACATATGATATCTTCTATGTTGTGAAA 1180
QY 1081 AGAGTTTTCGAAGATATGCAATTCGAATTTTGTAGTCCAGCCCTATATTTGCT 1140
Db 1181 AGAGTTTTCGAAGATATGCAATTCGAATTTTGTAGTCCAGCCCTATATTTGCT 1240
QY 1141 GGTGAGATATGCTGACTCATTTGAATCAAAAGCCATTTTGTGAATTTATGTTTTC 1200
Db 1241 GGTGAGATATGCTGACTCATTTGAATCAAAAGCCATTTTGTGAATTTATGTTTTC 1300
QY 1201 ATATGCTGTTTCACTGTTTATGTTTCTCAGAAAGCTGCTGGAATTTGTTACTTCA 1260
Db 1301 ATATGCTGTTTCACTGTTTATGTTTCTCAGAAAGCTGCTGGAATTTGTTACTTCA 1360
QY 1261 CCTTATGCTATTAATGAGGCTTAACATTAACCTGCTCCACATCCAGACTGTTTGTGA 1320
Db 1361 CCTTATGCTATTAATGAGGCTTAACATTAACCTGCTCCACATCCAGACTGTTTGTGA 1420
QY 1321 CTGAGTTCATGCAATTTGTTAATTTCAATCTTTTTCATCTTTTGTGAACAAGACTTT 1380
Db 1421 CTGAGTTCATGCAATTTGTTAATTTCAATCTTTTTCATCTTTTGTGAACAAGACTTT 1480
QY 1381 CAGTGGCCAATATGTCAGACATTCAAAGSTTATGTTGAATATCATGATATTTTGA 1440
Db 1481 CAGTGGCCAATATGTCAGACATTCAAAGSTTATGTTGAATATCATGATATTTTGA 1540
QY 1441 CTGTAAAAATGSACTTAATAA-TAGACATTTCTACAAAAGACACATGATAGGNGAAA 1499
Db 1541 CTGTAAAAATGSACTTAATAA-TAGACATTTCTACAAAAGACACATGATAGGNGAAA 1600
QY 1500 ACATGGAATTTCTTTAGTGCAGTGTGCTTCAAAATTTACATTTAGTTTAAATATA 1559
Db 1601 ACATGGAATTTCTTTAGTGCAGTGTGCTTCAAAATTTACATTTAGTTTAAATATA 1660
QY 1560 TATTTTAAACATATGTAGAAATTAAGTGCAAAAGACATGGAAGCTTAAAGCTGCTT 1619
Db 1661 TATTTTAAACATATGTAGAAATTAAGTGCAAAAGACATGGAAGCTTAAAGCTGCTT 1720

QY 1620 CAANAGCCTGAATTAATGGAAAAATAAAMNGTTTNCAGATATCTCAT 1667
Db 1721 CAAAAGCCTGAAAAATGAAAAATAAATTTGTTTCAGATATCTCAAA 1768

RESULT 4
AX591660
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

AX591660 1461 bp DNA linear PAT 27-JAN-2003
Sequence 21 from Patent WO0246409.
AX591660.1 GI:27950045
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.

REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
source
BASE COUNT
ORIGIN

Guo,X., Li,L., Patturajan,M., Shinkets,R.A., Casman,S.J.,
Malvanekar,U.M., Tchernev,V.T., Vernet,C.A., Spytek,K.A.,
Shenoy,S.G., Alsobrook,J.P., Edinger,S., Peyman,J.A., Stone,D.J.,
Ellerman,K., Gangolli,E.A., Bolodg,F.L., Colman,S.D., Eissen,A.J.,
Liu,X., Padigaru,M., Spaderna,S.K. and Zernhusen,B.D.
Proteins and nucleic acids encoding same
Patent: WO 0246409-A 21 13-JUN-2002;
Curagen Corporation (US)
Location/Qualifiers
1..1461
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

347 a 298 c 277 g 539 t

Query Match 73.2%; Score 1359.8; DB 6; Length 1461;
Best Local Similarly 97.7%; Pred. No. 4; 6e-282;
Matches 1379; Conservative 0; Mismatches 32; Indels 0; Gaps 0;

QY 1 ATGGCGAGCTAGAGGGTTACTGTTCTCGCGCCCTTGAGCTGTACCTTTTATGTC 60
Db 49 ATGGCGAGCTAGAGGGTTACTGTTCTCGCGCCCTTGAGCTGTACCTTTTATGTC 108
QY 121 CACCTGCTGAGCGCAGCGCTACTGTGAGGGCCATTTCTCCCTTCCAGTGGATCC 180
Db 169 CACCTGCTGAGCGCAGCGCTACTGTGAGGGCCATTTCTCCCTTCCAGTGGATCC 228
QY 181 ATGATTAATCACTTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 240
Db 229 ATGATTAATCACTTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 288
QY 241 TGGATCTTTGGATGCTGCAACATGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 300
Db 289 TGGATCTTTGGATGCTGCAACATGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 348
QY 301 CTTCCTCTCAAGTGTGGCAACTTCTATTTCTATATTTGCTTTTCCAGAGTCAACC 360
Db 349 CTTCCTCTCAAGTGTGGCAACTTCTATTTCTATATTTGCTTTTCCAGAGTCAACC 408
QY 361 AGAACAAGCGCTGCTCAACTATTCAGAGAGTCTGTGCAACATTAACACTAGAGATT 420
Db 409 AGAACAAGCGCTGCTCAACTATTCAGAGAGTCTGTGCAACATTAACACTAGAGATT 468
QY 421 CCAACACTTATTTTATTTTAACTTCTTATATATACAGAAACAGAGATCTATGTTTACT 480
Db 469 CCAACACTTATTTTATTTTAACTTCTTATATATACAGAAACAGAGATCTATGTTTACT 528
QY 481 CTTTTCATATTTGATGTCTTTTATGAAATCATATAAATCTTACGCTTCTTGGATT 540
Db 529 CTTTTCATATTTGATGTGTCTTTTATGAAATCATATAAATCTTACGCTTCTTGGATT 588

QY	541	TGTCCTTCATGTTTGGCAAAATATCATCTGGGCTGTCTTCTGTGACGGAATGTC	600
	589	TGTGCTTCATGTTTGGCAAAATATCATCTGGGCTGTCTTCTGTGACGGAATGTC	648
QY	601	ATTGCACAAAAGTTAACTGAGGCTTGGAAAACCTGACTACAAAAGGAAGACAGACTT	660
	649	ATTGCACAAAAGTTAACTGAGGCTTGGAAAACCTGACTACAAAAGGAAGACAGACTT	708
QY	661	CCACCTTTAAAGGACCATTTGACAAATTCAGAAAATTTCTGACTTTCTTTGGCTTAT	720
	709	CCACCTTTAAAGGACCATTTGACAAATTCAGAAAATTTCTGACTTTCTTTGGCTTAT	768
QY	721	TCCATGTCCTTTAAAAAATTGATGATCTTTCTGTTGACTTGGCCCTACATCTCTG	780
	769	TCCATGTCCTTTAAAAAATTGATGATCTTTCTGTTGACTTGGCCCTACATCTCTG	828
QY	781	GGATTCTGTTTGTGCTTTGTAGTATGTTATGTTGGAATGTTATTTGGCATGAGAT	840
	829	GGATTCTGTTTGTGCTTTGTAGTATGTTATGTTGGAATGTTATTTGGCATGAGAT	888
QY	841	AGTCATGAGGCTGTCTTCATTTTCCACACTATTTCTACTTTTCTATTTACTCTCTT	900
	889	AGTCATGAGGCTGTCTTCATTTTCCACACTATTTCTACTTTTCTATTTACTCTCTT	948
QY	901	TTTTCTTTCTCTCATCTCTGTCCTAGCAAAATTAAGACTTTTCTTCTTGTGTTGG	960
	949	TTTTCTTTCTCTCATCTCTGTCCTAGCAAAATTAAGACTTTTCTTCTTGTGTTGG	1008
QY	961	AAACATGGAATTCGTTTGTGTTGTTACCTTACTCTCTGTTTGTAGTTGGAATTC	1020
	1009	AAACATGGAATTCGTTTGTGTTGTTACCTTACTCTCTGTTTGTAGTTGGAATTC	1068
QY	1021	ACTATGCTCATTAATATCTGCTGACAGACAAATGACATTAATTTCTATGTTGAAA	1080
	1069	ACTATGCTCATTAATATCTGCTGACAGACAAATGACATTAATTTCTATGTTGAAA	1128
QY	1081	AGAGTTTTTCAAGATATGCAATTCGAAATTTTGTAGTCCAGGCTATATATTTGCT	1140
	1129	AGAGTTTTTCAAGATATGCAATTCGAAATTTTGTAGTCCAGGCTATATATTTGCT	1188
QY	1141	GGTTGGATATGCTGCTATTTGAATCAAGGCAATTTTGGATTTTAATGTTTTC	1200
	1189	GGTTGGATATGCTGCTATTTGAATCAAGGCAATTTTGGATTTTAATGTTTTC	1248
QY	1201	ATAATGCTTTCATTTTATAGTCTCTCAGAAAACCTGGAATTTCTTACTTCTATTTA	1260
	1249	ATAATGCTTTCATTTTATAGTCTCTCAGAAAACCTGGAATTTCTTACTTCTATTTA	1308
QY	1261	CCTTATGCTATTTATAGGCTTAACATACCTGCTCCACATCCAGACATTTGTTGAA	1320
	1309	CCTTATGCTATTTATAGGCTTAACATACCTGCTCCACATCCAGACATTTGTTGAA	1368
QY	1321	CTGAGTGCTATGCAATTTTAAATTTTAACTTTTAACTTTTAACTTTTAACTTTTAA	1380
	1369	CTGAGTGCTATGCAATTTTAAATTTTAACTTTTAACTTTTAACTTTTAACTTTTAA	1428
QY	1381	CAGTGGCCAAATAGTCAGACATTTCAAAGCT 1411	
	1429	CAGTGGCCAAATAGTCAGACATTTCAAAGCT 1459	

REFERENCE	1 (bases 1 to 165301)
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
	Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,
	Alshrocks,S.L., Amaraltinge,H.C., Are,J.R., Ayale,M., Banks,T.,
	Barbata,J., Benton,J., Bimaye,K., Blankenbury,K., Bonini,D.,
	Bouck,J., Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P.,
	Buhay,C., Burck,P., Burkett,C., Burrell,K.L., Byrd,N.C.,
	Carroll,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,
	Chen,G., Chen,R., Chen,Z., Chiu,D., Chokhry,I., Christopoulos,C.,
	Cleveland,C.D., Cox,C., Coyle,M.D., Dabonne,S.R., David,R.,
	Dayle,M.L., Davis,C., Davy-carroll, L., Deckerich,D.A.,
	Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Din,H.H.,
	Doutlawale,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,
	Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Emmerling,S.,
	Escotto,M., Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P.,
	Frantz,P., Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N.,
	Gill,R., Correll,J.H., Guevara,W., Gunaratne,P., Hale,S.,
	Hamilton,K., Han,J., Harris,C., Harris,K., Hart,M., Havlak,P.,
	Hawes,A., Hernandez,J., Hernandez,O., Hodgson,A., Hognes,M.,
	Holloway,C., Hollins,B., Homs,I.F., Howard,S., Huber,J., Hulyk,S.,
	Hume,J., Ioshikhes,I., Jackson,L.E., Jacobson,B., Jia,Y.,
	Johnson,R., Jolivet,S., Joudah,S., Karlsson,E., Kelly,S., Khan,U.,
	King,B., Korvah,J., Kovar,C., Kratovic,J., Kureshi,A., Landry,N.,
	Leal,B., Lee,E., Lewis,L.C., Lewis,L., Li,J., Li,Z., Lichteberg,O.,
	Lieu,C., Liu,J., Liu,W., Louised,H., Lozada,R.J., Lu,X.,
	Lucier,A., Lucier,R., Luna,R., Ma,J., Maheshwari,M., Mapa,P.,
	Marandel,I., Martin,R., Martindale,A., Martinez,E., Massey,E.,
	McWhiney,E., McLeod,M.P., Meador,M., Mei,G., Mercher,S.,
	Metzker,M., Miller,A., Miner,G., Miner,Z., Mitchell,T.,
	Monabdat,K., Montgomery,K.T., Morgan,M., Morris,S., Moser,M.,
	Neal,D., Nelson,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N.,
	Nguyen,N., Nickerson,E., Nwokenwo,S., Ogih,M., Okunou,G.,
	Oragunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L.,
	Peters,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y.,
	Rivers,M., Rojas,A., Rojibokan,I., Rolfe,M., Ruiz,S., Savery,G.,
	Schreier,S., Scott,G., Shen,H., Shim,C., Shoostari,N., Sisson,I.,
	Sodergren,E., Sonakke,T., Sparks,A., Stanley,H., Stone,H.,
	Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H.,
	Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N., Thomas,S.,
	Umanit,K., Vasquez,L., Vera,Y., Villalona,D., Vinson,R., Wang,Q.,
	Wang,S., Ward-Moore,S., Warren,R., Washington,C., Wellington,S.,
	Williams,G., Williamson,A., Wleczky,R., Wooden,S., Worley,K.,
	Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Kucherlapati,R.,
	Weinstock,G., and Gibbs,R.
TITLE	Direct Submission
JOURNAL	Unpublished
REFERENCE	2 (bases 1 to 165301)
AUTHORS	Morley,K.C.
TITLE	Direct Submission
JOURNAL	Submitted (10-APR-2002) Human Genome Sequencing Center, Department
	of Molecular and Human Genetics, Baylor College of Medicine, One
	Baylor Plaza, Houston, TX 77030, USA
REFERENCE	3 (bases 1 to 165301)
AUTHORS	Worley,K.C.
TITLE	Direct Submission
JOURNAL	Submitted (25-APR-2002) Human Genome Sequencing Center, Department
	of Molecular and Human Genetics, Baylor College of Medicine, One
	Baylor Plaza, Houston, TX 77030, USA
REFERENCE	4 (bases 1 to 165301)
AUTHORS	Worley,K.C.
TITLE	Direct Submission
JOURNAL	Submitted (30-JUL-2002) Human Genome Sequencing Center, Department
	of Molecular and Human Genetics, Baylor College of Medicine, One
	Baylor Plaza, Houston, TX 77030, USA
COMMENT	On Apr 25, 2002 this sequence version replaced g1:20127350.
	INFORMATION: http://www.hgsc.bcm.tmc.edu/ or email
	gc-help@bcm.tmc.edu

CLONE LENGTH: This sequence does not necessarily represent the entire insert of this clone. Overlapping regions of clones are only sequenced and submitted once, so the sequence for the remainder of the insert may be found in the record for the adjacent clones.

Features listing.

ANNOTATION OF FEATURES:

STSs are identified using ePCR (Genome Res. 7:541-550) searches of a local database that includes entries from dbSTS, GDB, and local mapping efforts.

Repeats are identified using RepeatMasker (A. Smit and P. Green, unpublished.) for Human and Mouse sequences. Genes and Region of sequence similarity are identified by BLAST (Nuc. Acids Res. 25:3389-3402) similarity (expect < 1e-34) to the EST and cDNA sequences. Genes demonstrate at least two exons flanked by consensus splice sites that maintained sequence continuity across the splice junctions. Sequences that are not identical matches are annotated as similar.

SEQUENCING READ COVERAGE: Sequencing is completed to a minimum standard of double strand coverage with a minimum of 2 clones and 2 reads with no ambiguities or 2 chemistries with a minimum of 2 clones and 3 reads with no ambiguities. If the sequence quality for a region does not meet this standard, it will be indicated in the annotation as Low Coverage.

QUALITY OF INDIVIDUAL BASES: This sequence meets stringent quality standards - estimated error rate less than 1 per 10,000 bases. Reports of lowest quality individual bases and measures of base quality are listed below. Description of the metrics can be found at URL: <http://gc.bom.tmc.edu:8088/quality.info/genbank.annotation.html>.

QUALSTAT-REPORT.

FEATURES
source Location/Qualifiers

```
1. .165301
  /organism="Homo sapiens"
  /mol_type="genomic DNA"
  /db_xref="taxon:9606"
  /chromosome="12"
  /clone="RP11-266015"
  1. .1600
  /note="overlaps bases 174395. .175994 of clone AC079460"
  /function="clone overlap"
  repeat_region
    /rpt_family="AT_rich"
    124. .144
  repeat_region
    /rpt_family="AT_rich"
    271. .304
  repeat_region
    /rpt_family="AT_rich"
    305. .829
  repeat_region
    /rpt_family="LIPAl6"
    846. .1506
  repeat_region
    /rpt_family="LIM1"
    1460. .1602
  repeat_region
    /rpt_family="LIM1"
    1586. .3715
  repeat_region
    /rpt_family="LIM1"
    3716. .3917
  repeat_region
    /rpt_family="LIPAl2"
    3918. .5682
  repeat_region
    /rpt_family="LIM1"
    5702. .5973
  repeat_region
    /rpt_family="LIPAl5"
    5974. .6253
  repeat_region
    /rpt_family="AluXs"
    6257. .6313
  repeat_region
    /rpt_family="(CAAAA)n"
    6314. .6684
  repeat_region
    /rpt_family="LIPAl5"
    6685. .7103
  repeat_region
    /rpt_family="MSTAl"
    7104. .7299
  repeat_region
    /rpt_family="MSTAl-internal"
    7303. .8605
  repeat_region
    /rpt_family="MSTAl-internal"
    8608. .9012
  repeat_region
    /rpt_family="MSTAl"
    9013. .10019
```

```

/rpt_family="LIPAl5"
repeat_region
  10026. .10345
/rpt_family="THE1C"
repeat_region
  10346. .10631
/rpt_family="AluSg1"
repeat_region
  10632. .10691
/rpt_family="THE1C"
repeat_region
  10720. .10829
/rpt_family="LIMAl2"
repeat_region
  10830. .10851
/rpt_family="AT_rich"
repeat_region
  complement(10852. .10976)
/rpt_family="MER4B"
repeat_region
  11307. .11588
/rpt_family="AluXs"
repeat_region
  11590. .11626
/rpt_family="AT_rich"
repeat_region
  12233. .12406
/rpt_family="MER41D"
repeat_region
  13756. .13929
/rpt_family="MER66-internal"
repeat_region
  14174. .14674
/rpt_family="MER66-internal"
repeat_region
  complement(15308. .15623)
/rpt_family="AluY"
repeat_region
  15758. .16083
/rpt_family="HERV17"
repeat_region
  16551. .16682
/rpt_family="Harleguin"
repeat_region
  complement(16881. .17089)
/rpt_family="THE1B"
repeat_region
  17090. .17350
/rpt_family="AluXs"
repeat_region
  17351. .17373
/rpt_family="(TAA)n"
repeat_region
  complement(17375. .17537)
/rpt_family="THE1B"
repeat_region
  17695. .17992
/rpt_family="HUIERS-P3"
repeat_region
  18396. .18682
/rpt_family="HERV17"
repeat_region
  19410. .19466
/rpt_family="G-rich"
repeat_region
  24062. .24174
/rpt_family="(CA)n"

Query Match      72.7%; Score 1349.2; DB 9; Length 165301;
Best Local Similarity 97.6%; Pred. No. 8.4e-280;
Matches 1356; Conservative 8; Mismatches 24; Indels 1; Gaps 1;

QY      366  CAAGGCTGGCTCAAGATATCCAGAGAGCTTTGTCAACATTTACACTGACATATTTCCAC 425
DB      137761 CAAGGCTGGCTCAAGATATCCAGAGAGCTTTGTCAACATTTACACTGACATATTTCCAC 137820
QY      426  ACTTTATTTTAACTCCCTTATATATACAGAGAGATCTATGTTTAACTCTTT 485
DB      137821 ACTTTATTTTAACTCCCTTATATATACAGAGAGATCTATGTTTAACTCTTT 137880
QY      486  TGCATATTTGATGTGCTTATGGAATATATAAACTTCAGCCCTTGATTTTGTG 545
DB      137881 TGCATATTTGATGTGCTTATGGAATATATAAACTTCAGCCCTTGATTTTGTG 137940
QY      546  CTTTCATGTTGGGCAACAATATATATCTGGGTCTCTCTGTCAGGGAATGTCAT 605
DB      137941 CTTTCATGTTGGGCAACAATATATCTGGGTCTCTCTGTCAGGGAATGTCAT 138000
QY      606  ACAAAAGTTAACGAGGCTTGGAAACTAGCTACAAAAGAGAGACAGACTTCACC 665
DB      138001 ACAAAAGTTAACGAGGCTTGGAAACTAGCTACAAAAGAGAGACAGACTTCACC 138060
QY      666  TATTAAAGACCAATTTGACGAATTCAGAAAATCTTCAGTTCTTTGGCTTATTCAT 725
DB      138061 TATTAAAGACCAATTTGACGAATTCAGAAAATCTTCAGTTCTTTGGCTTATTCAT 138120
```

OY	726	GTGCTTTAAAAACCTGAGCTATGCTTTGCTTTGAGCTTGGCCCTACACCTCTGGAGT	785
Db	138121	GTGCTTTAAAAACCTGAGCTATGCTTTGCTTTGAGCTTGGCCCTACACCTCTGGAGT	138180
OY	786	TCTGCTTTTGTGCTTTTGTAGTAGTTAAAGTGCGAATTTGTATTGGCGAGTCGAGTAGTCA	845
Db	138181	TCTGCTTTTGTGCTTTTGTAGTAGTTAAAGTGCGAATTTGTATTGGCGAGTCGAGTAGTCA	138240
OY	846	TGAAGCCGTGCTTCATTTTCTGCAACATTTCTACTTTTTCATTACTTCTCTTTTTC	905
Db	138241	TGAAGCCGTGCTTCATTTTCTGCAACATTTCTACTTTTTCATTACTTCTCTTTTTC	138300
OY	906	TTTTCGTCATCTCCGCTCCTCAGCAAAATTAAGACTTTTCTCCTATGTTTGGAAACA	965
Db	138301	TTTTCGTCATCTCCGCTCCTCAGCAAAATTAAGACTTTTCTCCTATGTTTGGAAACA	138360
OY	966	TGGAATCTGTTTTTGGTGCTTACCTTAGTCTGTGCTTTTGTAGTTTGGAAATTCCTTA	1025
Db	138361	TGGAATCTGTTTTTGGTGCTTACCTTAGTCTGTGCTTTTGTAGTTTGGAAATTCCTTA	138420
OY	1026	TGCGCATTAATACCTTGTAGCGACGAACAATAGACATTAATCTTCTATGTTGGGAAAGAGT	1085
Db	138421	TGCGCATTAATACCTTGTAGCGACGAACAATAGACATTAATCTTCTATGTTGGGAAAGAGT	138480
OY	1086	TTTTCAAGATATGCAATTTTGAAATTTTGTAGTTTCCAGCCTATATTTTGTGCTGTG	1145
Db	138481	TTTTCAAGATATGCAATTTTGAAATTTTGTAGTTTCCAGCCTATATTTTGTGCTGTG	138540
OY	1146	GAGTATACCTACACATGTAAGAAATCAAGGCATTTTGGAAATTAATCTTTTTCATAG	1205
Db	138541	GAGTATACCTACACATGTAAGAAATCAAGGCATTTTGGAAATTAATCTTTTTCATAG	138600
OY	1206	CTTGCTTCAATGTTATAGTTTCCACAGAACTGCTGGAATTTGCTTACTTCAATTTACCTTA	1265
Db	138601	CTTGCTTCAATGTTATAGTTTCCACAGAACTGCTGGAATTTGCTTACTTCAATTTACCTTA	138660
OY	1266	TGTCAATTTATAGGCTTAACATATACCTGCTCCACATCCAGACTGTTTGTGAACCTGAG	1325
Db	138661	TGTCAATTTATAGGCTTAACATATACCTGCTCCACATCCAGACTGTTTGTGAACCTGAG	138720
OY	1326	TTGCGATGCAATTTGTTAATTTTCATTAACCTTTTACATCTTTCGAAACAACACTTTTCAGT	1385
Db	138721	TTGCGATGCAATTTGTTAATTTTCATTAACCTTTTACATCTTTCGAAACAACACTTTTCAGT	138780
OY	1386	GCCAAATATGTCAGGACATTTCAAAAGGTTTATGTGTATATCAAGTAGATTTTGTGAAGCTTA	1445
Db	138781	GCCAAATATGTCAGGACATTTCAAAAGGTTTATGTGTATATCAAGTAGATTTTGTGAAGCTTA	138840
OY	1446	AAAAATGAGCTTAATATA-TAGACCATTTTTCACAAAGAACACAGTAAGGCGAAAAATG	1504
Db	138841	AAAAATGAGCTTAATATAATTAGACCATTTTTCACAAAGAACACAGTAAGGCGAAAAATG	138900
OY	1505	GAAATTTCTTTAGGTGACAGTGTGCTTCAATTAACATTAAGTTTAAAAATATATATTT	1564
Db	138901	GAAATTTCTTTAGGTGACAGTGTGCTTCAATTAACATTAAGTTTAAAAATATATATTT	138960
OY	1565	TAAACATATGTAAGAAATTAAGTGGCGAAAGACTGGGAAAGCTTAGACTGCTTCAAN	1624
Db	138961	TAAACATATATGTAAGAAATTAAGTGGCGAAAGACTGGGAAAGCTTAGACTGCTTCAAN	139020
OY	1625	GCGTCAATAATAGGAAAAATTAANMNGTTTNCAGATATCCATATCCCTCINNANATAGT	1684
Db	139021	GCGTCAATAATAGGAAAAATTAANMNGTTTTCAGATATCCATATCCCTCINNANATAGT	139080
OY	1685	GCCCTTMCANANGCTTGGGAATGKTTNNNTGNATATAGTTNATTAANCTGGGNNTGCTN	1744
Db	139081	GCCCTTMCANANGCTTGGGAATGTTTGTATGTATATAGTTTATTAACCTGGGTATAGCTT	139140
OY	1745	NNNATNACT 1753	
Db	139141	CATATTACT 139149	

RESULT 6	AC024000/c	179032 bp	DNA	linear	HTG 08-MAR-2000
LOCUS	AC024000	179032 bp	DNA	linear	HTG 08-MAR-2000
DEFINITION	Homo sapiens chromosome 6 clone RP11-266015 map 6, WORKING DRAFT				
SEQUENCE	SEQUENCE, 35 unordered pieces.				
AC024000					
AC024000.2	GI:7210053				
AC024000.2	GI:7210053				
KEYWORDS	HTG; HTGS_PHASE1; HTGS_DRAFT.				
SOURCE	Homo sapiens				
ORGANISM	Homo sapiens (human)				
REFERENCE	Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.				
AUTHORS	1 (bases 1 to 179032)				
JOURNAL	Birren, B., Linton, L., Nusbaum, C. and Lander, E.				
TITLE	Homo sapiens chromosome 6, clone RP11-266015				
AUTHORS	Unpublished				
AUTHORS	2 (bases 1 to 179032)				
AUTHORS	Anderson, S., Baldwin, J., Barna, N., Bede, F., Boguslavsky, L., Boukhalter, B., Brown, A., Burkett, G., Campolano, A., Castle, A., Choepel, Y., Colangelo, M., Collins, S., Collymore, A., Cooke, P., Deaerellano, K., Dewar, K., Dodge, S., Domino, M., Doyle, M., Fenesor, J., Ferreira, P., FitzHugh, W., Forrest, C., Gage, D., Galagan, J., Gardyna, S., Ginde, S., Goyette, M., Graham, L., Howard-Plat, N., Grant, G., Hagos, B., Heaford, A., Horton, L., Howland, J. C., Iliev, I., Johnson, R., Jones, C., Kann, L., Karitas, A., Klein, J., Landers, T., Laroque, K., Lehotzky, J., Levine, R., Liu, C., Liu, G., Locke, K., MacDonald, P., Margulis, N., McCarthy, M., McEwan, P., McGurk, A., McKernan, K., McPheters, R., Meldrum, J., Menus, L., Miho, T., Miranda, C., Miengo, V., Morrow, J., Naylor, J., Norman, C. H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, T. M., Peterson, K., Pierre, N., Pisani, C., Pollara, V., Raymond, C., Riley, R., Rogov, P., Rothman, D., Roy, A., Santos, R., Schauer, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N., Sudriaman, A., Talamas, J., Testate, S., Theodore, J., Tirrell, A., Travers, M., Trigilio, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G., Zainoun, J., Zimmer, A. and Zody, M.				
TITLE	Direct Submission				
JOURNAL	Submitted (20-FEB-2000) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA				
COMMENT	On Mar 8, 2000 this sequence version replaced gi:7008905. All repeats were identified using RepeatMasker: Smit, A.F.A. & Green, P. (1996-1997) http://ftp.genome.washington.edu/RM/RepeatMasker.html				
	----- Genome Center				
	Center: Whitehead Institute/MIT Center for Genome Research				
	Center code: WIBR				
	Web site: http://www-seq.wi.mit.edu				
	Contact: sequence_submissions@genome.wi.mit.edu				
	----- Project Information				
	Center project name: L7057				
	Center clone name: 266_O_15				
	----- Summary Statistics				
	Sequencing vector: M13; M77815; 100% of reads				
	Chemistry: Dye-terminator Big Dye; 100% of reads				
	Assembly program: Phrap; version 0.960731				
	Consensus quality: 153613 bases at least Q40				
	Consensus quality: 164872 bases at least Q30				
	Consensus quality: 171071 bases at least Q20				
	Insert size: 175632; sum-of-coverage				
	Quality coverage: 3.8 in Q20 bases; sum-of-coverage				

	* NOTE: This is a 'working draft' sequence. It currently				
	* consists of 35 contigs. The true order of the pieces				
	* is not known and their order in this sequence record is				
	* arbitrary. Gaps between the contigs are represented as				
	* runs of N, but the exact sizes of the gaps are unknown.				
	* This record will be updated with the finished sequence				
	* as soon as it is available and the accession number will				
	* be preserved.				
	I 1242: contig of 1242 bp in length				

1243 1342: gap of 100 bp
1343 2622: contig of 1280 bp in length
2623 2722: gap of 100 bp
2723 3815: contig of 1093 bp in length
3816 3915: gap of 100 bp
3916 5846: contig of 1931 bp in length
5847 5946: gap of 100 bp
5947 8052: contig of 2106 bp in length
8053 8152: gap of 100 bp
8153 10003: contig of 1851 bp in length
10004 10103: gap of 100 bp
10104 13638: contig of 3535 bp in length
13639 13738: gap of 100 bp
13739 16505: contig of 2767 bp in length
16506 16605: gap of 100 bp
16606 19624: contig of 3019 bp in length
19625 19724: gap of 100 bp
19725 22156: contig of 2432 bp in length
22157 22256: gap of 100 bp
22257 24548: contig of 2292 bp in length
24549 24648: gap of 100 bp
24649 27929: contig of 3281 bp in length
27930 28029: gap of 100 bp
28030 30657: contig of 2628 bp in length
30658 30757: gap of 100 bp
30758 34014: contig of 3257 bp in length
34015 34114: gap of 100 bp
34115 35085: contig of 971 bp in length
35086 35185: gap of 100 bp
35186 39055: contig of 3870 bp in length
39056 39155: gap of 100 bp
39156 42847: contig of 3692 bp in length
42848 42947: gap of 100 bp
42948 47517: contig of 4570 bp in length
47518 47617: gap of 100 bp
47618 50774: contig of 3157 bp in length
50775 50874: gap of 100 bp
50875 54256: contig of 3382 bp in length
54257 54356: gap of 100 bp
54357 58391: contig of 4035 bp in length
58392 58491: gap of 100 bp
58492 63149: contig of 4658 bp in length
63150 63249: gap of 100 bp
63250 68843: contig of 5594 bp in length
68844 68943: gap of 100 bp
68944 74638: contig of 5695 bp in length
74639 74738: gap of 100 bp
74739 80805: contig of 6067 bp in length
80806 80905: gap of 100 bp
80906 86709: contig of 5804 bp in length
86710 86809: gap of 100 bp
86810 93296: contig of 6487 bp in length
93297 93396: gap of 100 bp
93397 100379: contig of 6983 bp in length
100380 100479: gap of 100 bp
100480 110208: contig of 9729 bp in length
110209 110308: gap of 100 bp
110309 118705: contig of 8397 bp in length
118706 118805: gap of 100 bp
118806 130465: contig of 11660 bp in length
130466 130565: gap of 100 bp
130566 141090: contig of 10525 bp in length
141091 141190: gap of 100 bp
141191 154350: contig of 13160 bp in length
154351 154450: gap of 100 bp
154451 166361: contig of 11911 bp in length
166362 179032: gap of 100 bp
179033 179032: contig of 12571 bp in length.

FEATURES
source

1. 179032
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

/chromosome="6"
/map="6"
/clone="RP11-266015"
/clone_11b="RPC1-11 Human Male BAC"
1. 1242
/note="assembly_fragment"
1343. 2622
/note="assembly_fragment"
2723. 3815
/note="assembly_fragment"
3916. 5846
/note="assembly_fragment"
5947. 8052
/note="assembly_fragment"
8153. 10003
/note="assembly_fragment"
10104. 13638
/note="assembly_fragment"
13739. 16505
/note="assembly_fragment"
16606. 19624
/note="assembly_fragment"
19725. 22156
/note="assembly_fragment"
22257. 24548
/note="assembly_fragment"
24649. 27929
/note="assembly_fragment"
28030. 30657
/note="assembly_fragment"
30758. 34014
/note="assembly_fragment"
clone_end:T7
vector_side:right"
34115. 35085
/note="assembly_fragment"
clone_end:SP6
vector_side:right"
35186. 39055
/note="assembly_fragment"
39156. 42847
/note="assembly_fragment"
42948. 47517
/note="assembly_fragment"
47618. 50774
/note="assembly_fragment"
50875. 54256
/note="assembly_fragment"
54357. 58391
/note="assembly_fragment"
58492. 63149
/note="assembly_fragment"
63250. 68843
/note="assembly_fragment"
68944. 74638
/note="assembly_fragment"
74739. 80805
/note="assembly_fragment"

Query Match 72.7%; Score 1349.2; DB 2; Length 179032;
Best Local Similarity 97.6%; Pred. No. 8.4e-280;
Matches 1356; Conservative 8; Mismatches 24; Indels 1; Gaps 1;

QY 366 CAAGGCTGCCCTCAAGATCCAGAGAGTCTTGTATACATTTACACTGACGATTTTCCAC 425
DB 9587 CAAGGCTGCCCTCAAGATCCAGAGAGTCTTGTATACATTTACACTGACGATTTTCCAC 9528
QY 426 ACTTATTTTAACTCCCTTATATACAGAGAGATCATGTTTTCCTCTTT 485
DB 9527 ACTTATTTTAACTCCCTTATATACAGAGAGATCATGTTTTCCTCTTT 9468
QY 486 TGCATATTTGATGTCTTTATGGAATACATAAACTTCAGCCTTCCTGATTTTGTGG 545
DB 9467 TGCATATTTGATGTCTTTATGGAATACATAAACTTCAGCCTTCCTGATTTTGTGG 9408

QY	546	CTTCATGCTTTGGCAAAACAATATCATCTGGGCTGTCTCTGTGTCAGGGAATGTCAATTCG	605
Dd	9407	CTTCATGCTTTGGCAAAACAATATCATCTGGGCTGTCTCTGTGTCAGGGAATGTCAATTCG	9348
QY	606	ACAAAAGTTAACTGAGCGCTTGAAAAACAGCACTACAAAAGAGAAGACAGACTTCCACC	665
Dd	9347	ACAAAAGTTAACTGAGCGCTTGAAAAACAGCACTACAAAAGAGAAGACAGACTTCCACC	9288
QY	666	TATTAAAGACCACTTTGCAGAAATTCAGAAAATCTTCAGTTCTCTTTGGCTTATTCAT	725
Dd	9287	TATTAAAGACCACTTTGCAGAAATTCAGAAAATCTTCAGTTCTCTTTGGCTTATTCAT	9228
QY	726	GTCCCTTTAAAAACGTGAGTATGACCTTTCTGTTGACCTGGCCACATCCCTTGGGAGTT	785
Dd	9167	TCGTGTTTGTCTTTTGAGTAGTTAATAGGGAATTTGTATTTGGCGATCGGAGTAGTCA	9108
QY	846	TGAAGCCCTGCTCATTTTCTCTCAACTTCTACTTTTTCATTATTACTCTCTTTTTC	905
Dd	9107	TGAAGCCCTGCTCATTTTCTCTCAACTTCTACTTTTTCATTATTACTCTCTTTTTC	9048
QY	906	TTTTCCTCATCTCTGTCTCTCTAGCAAAATTAAGACTTCTTCTTCCCTAGTTGGAACA	965
Dd	9047	TTTTCCTCATCTCTGTCTCTCTAGCAAAATTAAGACTTCTTCTTCCCTAGTTGGAACA	8988
QY	966	TGGAATTCGTCTTTTGGTGGTTACCTTAGTCTGTGTGTTTTAGTTTGGAAATTCACTTA	1025
Dd	8987	TGGAATTCGTCTTTTGGTGGTTACCTTAGTCTGTGTGTTTTAGTTTGGAAATTCACTTA	8928
QY	1026	TGCTCATTAATACCTGTGACAGACAATAAGACATTATCTTCTATGTGGAAGAAAGT	1085
Dd	8927	TGCTCATTAATACCTGTGACAGACAATAAGACATTATCTTCTATGTGGAAGAAAGT	8868
QY	1086	TTTTCAAGATATGCAATTCGAAATATTTGTAGTTCCAGCCATATATTTTGGCTGTG	1145
Dd	8867	TTTTCAAGATATGCAATTCGAAATATTTGTAGTTCCAGCCATATATTTTGGCTGTG	8808
QY	1146	GAGTATATGCTACATCTGAAATCAAAAGCCATTTTGGAAATTAAGTTTTCATATG	1205
Dd	8807	GAGTATATGCTACATCTGAAATCAAAAGCCATTTTGGAAATTAAGTTTTCATATG	8748
QY	1206	CTTGTTCATCTTATAGTTCCTCGAAACCTCTGGAATTTGTTACTTCATTTTACCTTA	1265
Dd	8747	CTTGTTCATCTTATAGTTCCTCGAAACCTCTGGAATTTGTTACTTCATTTTACCTTA	8688
QY	1266	TGTCAATTATAGGCTTAAACATTAACCTGTGCTCCACATCCAGACTTGTTTGTGAACGTAG	1325
Dd	8687	TGTCAATTATAGGCTTAAACATTAACCTGTGCTCCACATCCAGACTTGTTTGTGAACGTAG	8628
QY	1326	TTTGCATGCAATTTGTTAAATTCATTAACCTTTTAACTTTCTTGAAACAAGACTTTTCAGTG	1385
Dd	8627	TTTGCATGCAATTTGTTAAATTCATTAACCTTTTAACTTTCTTGAAACAAGACTTTTCAGTG	8568
QY	1386	GCCAAATATGTCAGACATTCAAAGCTTATGTGTAATATACGTATATTTTGAACGTGA	1445
Dd	8567	GCCAAATATGTCAGACATTCAAAGCTTATGTGTAATATACGTATATTTTGAACGTGA	8508
QY	1446	AAAATGGCACTTAATAA-TAGACCACTTCTACAAAAGAACACTGAAATGAGNGGAAAACATG	1504
Dd	8507	AAAATGGCACTTAATAA-TAGACCACTTCTACAAAAGAACACTGAAATGAGNGGAAAACATG	8448
QY	1505	GAAATTTCTTTTAGGTGACAGTGTGTGCTTCAAAATTTACATTTAGTTTTTTATATATATTT	1564
Dd	8447	GAAATTTCTTTTAGGTGACAGTGTGTGCTTCAAAATTTACATTTAGTTTTTTATATATATTT	8388
QY	1565	TAAACATATATGTAAGAAATTAAGTGGCAAGAACTGGGAAGCTTAAGACCTGCTTCAAA	1624
Dd	8387	TAAACATATATGTAAGAAATTAAGTGGCAAGAACTGGGAAGCTTAAAGCCGTCTTCAAA	8328

QY	1625	GCCTAAATAAGGGGAAATTAANNGGCTTTCGAGATATCCAAATGCTGCTC NNKNNKATGTCG	1684
Db	8327	GCCTAAATAAGGGGAAATTAANNGGCTTTCGAGATATCCAAATGCTGCTC NNKNNKATGTCG	8266
QY	1685	GCCCTTMCANAGCTTGGGAATGKTTNNMNTGNATTAAGTTAATTAANCCTGGGNNTCCTN	1744
Db	8267	GCCCTTMCANAGCTTGGGAATGKTTNNMNTGNATTAAGTTAATTAANCCTGGGNNTCCTN	8208
QY	1745	NNMATNACT 1753	
Db	8207	CATATTACT 8199	
RESULT 7			
AX591658			
LOCUS	AX591658	1489 bp	DNA linear PAT 27-JAN-2003
DEFINITION	Sequence 19 from Patent WO0246409.		
ACCESSION	AX591658		
VERSION	AX591658.1	GI:27950044	
KEYWORDS			
SOURCE	Homio sapiens (human)		
ORGANISM	Homio sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.		
AUTHORS	Guo, X., Li, L., Patturajan, M., Shlunkets, R.A., Casman, S.J., Malyanar, U.W., Tchernev, V.T., Verne, C.A., Spytek, K.A., Shenoy, S.G., Alsbrook, J.P., Edinger, S., Peyman, J.A., Stone, D.J., Ellerman, K., Gangoli, E.A., Boldog, F.L., Colman, S.D., Eisen, A.J., Liu, X., Padigaru, M., Spaderna, S.K., and Zernusen, B.D. Proteins and nucleic acids encoding same Patent: WO 0246409-A 19 13-JUN-2002; Curagen Corporation (US)		
TITLE			
JOURNAL			
FEATURES			
source	1. 1489 Location/Qualifiers		
	/organism="Homo sapiens"		
	/mol_type="genomic DNA"		
	/db_xref="taxon:9606"		
BASE COUNT	357 a 306 c 279 g 547 t		
ORIGIN			
Query Match	70.6%; Score 1311.8; DB 6; Length 1489;		
Best Local Similarity	95.4%; Pred. No. 9.9e-272;		
Matches 1374; Conservative	0; Mismatches 37; Indels 30; Gaps 1		
QY	1	ATGGCGCAGCTGAGAGGTTACTGTTTCTGCGCCGCTTGAGCTGTACTTTTAAGTTC	60
Db	49	ATGGCGCAGCTGAGAGGTTACTGTTTCTGCGCCGCTTGAGCTGTACTTTTAAGTTC	108
QY	61	TGCCCTCTTCTCCGCTTCAGCGGGGGGCTGCAGAGACCCTACATGAGAGATCTTC	120
Db	109	TGCCCTCTTCTCCGCTTCAGCGGGGGGCTGCAGAGACCCTACATGAGAGATCTTC	168
QY	121	CACCTGCCCTCAGCGCAGCGCTACTGTGAGGGCCATTTCTCCCTTCCCATGAGGATCC	180
Db	169	CACCTGCCCTCAGCGCAGCGCTACTGTGAGGGCCATTTCTCCCTTCCCATGAGGATCC	228
QY	181	ATGATTACTACATTACCTGCGCTTGACTCGTGTGATGAGTGGTCAAACTGCCATT	240
Db	229	ATGATTACTACATTACCTGCGCTTGACTCGTGTGATGAGTGGTCAAACTGCCATT	288
QY	241	TGATCTTGGATGGTGTGACATTTGTCTGCTCATTGGATGTCAGATTTGTAAAT	300
Db	289	TGATCTTGGATGGTGTGACATTTGTCTGCTCATTGGATGTCAGATTTGTAAAT	348
QY	301	CTTCTCTTCAGTGTGGCACTTCTATTACTATATTTGCTTTCCACAGGTACACCC	360
Db	349	CTTCTCTTCAGTGTGGCACTTCTATTACTATATTTGCTTTCCACAGGTACACCC	408
QY	361	AGAAACAAG-----GCTGCTCAAGTATCCAGAGA 390	
Db	409	AGAAACAAGTATTTTCAAAATACATTAATACAGTTGGCTGCTCAGATATCCAGAGA 468	


```

QY 391 GTCCTGCAACATTAAACAGTAGATTTTCCAACTATTTTAACTTCCTTAT 450
    |||||
Db 469 GTCCTGCAACATTAAACAGTAGATTTTCCAACTATTTTAACTTCCTTAT 528
QY 451 TATACAGACAGAGATCTATGTTTCTTCTTTTGCATATTTGATGTCTTTATGA 510
    |||||
Db 529 TATACAGACAGAGATCTATGTTTCTTCTTTTGCATATTTGATGTCTTTATGA 588
QY 511 AATCATAAACCTCAGCCTTCCTTGGATTTTGGGCTTCATGTTTGGGCAACAAATATC 570
    |||||
Db 589 AATCATAAACCTCAGCCTTCCTTGGATTTTGGGCTTCATGTTTGGGCAACAAATATC 648
QY 571 ATCTGGGCTGCTCTTCTGTCAGGAAATGTCATTCACAAAGTAACTAGAGCTTGAAA 630
    |||||
Db 649 ATCTGGGCTGCTCTTCTGTCAGGAAATGTCATTCACAAAGTAACTAGAGCTTGAAA 708
QY 631 ACTGAGCTACAAAGAAGAGACAGACTTCCACTATTAAGGACATTTGAGAAATTC 690
    |||||
Db 709 ACTGAGCTACAAAGAAGAGACAGACTTCCACTATTAAGGACATTTGAGAAATTC 768
QY 691 AGAAAAATCTCAGTTCTTTGGCTTATTCATGTCCTTAAAACTGAGTATGCT 750
    |||||
Db 769 AGAAAAATCTCAGTTCTTTGGCTTATTCATGTCCTTAAAACTGAGTATGCT 828
QY 751 TTCTGTTGACTTGGCCCTACATCTTCTGGGATTTCTGTTTGGCTTTGATAGTT 810
    |||||
Db 829 TTCTGTTGACTTGGCCCTACATCTTCTGGGATTTCTGTTTGGCTTTGATAGTT 888
QY 811 AATGCTGGAATTTTATGCGCATTCGAGTAGTCATGACCCGTCTTCATTTTCTCAA 870
    |||||
Db 889 AATGCTGGAATTTTATGCGCATTCGAGTAGTCATGACCCGTCTTCATTTTCTCAA 948
QY 871 CTAATCTACTTTTTCATTTACCTCTTTTCTTTTCTTCATCTCCCTGCTCCAGC 930
    |||||
Db 949 CTAATCTACTTTTTCATTTACCTCTTTTCTTTTCTTCATCTCCCTGCTCCAGC 1008
QY 931 AAAATTAAGACTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 990
    |||||
Db 1009 AAAATTAAGACTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1068
QY 991 TTATGCTCTGCTTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1050
    |||||
Db 1069 TTATGCTCTGCTTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1128
QY 1051 AATAGACATTAATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1110
    |||||
Db 1129 AATAGACATTAATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1188
QY 1111 TATTTGTTAGTCCAGCCTATATATTTGCTGTTGGAGTAGCTGACTGATTTGAATCA 1170
    |||||
Db 1189 TATTTGTTAGTCCAGCCTATATATTTGCTGTTGGAGTAGCTGACTGATTTGAATCA 1248
QY 1171 AAGCAATTTTGGAAATTAATGTTTCTCATATGCTTGTTCATTTGTAATGTTCTGAG 1230
    |||||
Db 1249 AAGCAATTTTGGAAATTAATGTTTCTCATATGCTTGTTCATTTGTAATGTTCTGAG 1308
QY 1231 AACTGCTGGAATTTGCTTACTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1290
    |||||
Db 1309 AACTGCTGGAATTTGCTTACTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1368
QY 1291 CTGCTCTCCACATCCAGACTTGTGTGAACTGAGTGTGATGCAATTTGTAATTTGATA 1350
    |||||
Db 1369 CTGCTCTCCACATCCAGACTTGTGTGAACTGAGTGTGATGCAATTTGTAATTTGATA 1428
QY 1351 ACTTTTACATCTTTTCTGAACAAGACTTTTCACTGGCCAAATATGTCAGACATTTCAAAG 1410
    |||||
Db 1429 ACTTTTACATCTTTTCTGAACAAGACTTTTCACTGGCCAAATATGTCAGACATTTCAAAG 1488
QY 1411 T 1411
    |
Db 1489 T 1489

```

RESULT 8
 AC046130
 LOCUS
 DEFINITION
 Homo sapiens 12 BAC RP11-847H18 (Roswell Park Cancer Institute Human BAC Library) complete sequence.
 AC046130
 AC046130.25 GI:11496301
 HTG.
 Homo sapiens (human)
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 1 (bases 1 to 176600)
 Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-Osman,F.R., Allen,C., Alstbrooks,S.L., Amaralunge,H.C., Ale,J.R., Banks,T.R., Barbatia,J., Benton,U., Bimaye,K., Blankenhorn,K., Bonnin,D., Bouck,J., Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P., Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C., Carton,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Z., Chiu,D., Chowdhury,I., Christopoulos,C., Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R., Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A., Delaney,K.R., Delgado,O., Dean,A.L., Ding,Y., Dinh,H.H., Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J., Earnhardt,C., Edgar,D., Edwards,C.C., Elhaj,C., Emeling,S., Escotto,M., Falls,T., Ferraguto,D., Flagg,A., Ford,J., Foster,P., Frantz,P., Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R., Gorrell,J.H., Guevara,M., Gunaratne,P., Hale,S., Hamilton,K., Han,J., Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., Hernandez,J., Hernandez,O., Hodgson,A., Hognes,M., Holloway,C., Hollins,B., Homai,F., Howard,S., Huber,J., Huliy,S., Hume,J., Ioshkhes,I., Jackson,L.E., Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S., Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C., Kratochvic,J., Kureshi,A., Landry,N., Leal,B., Lee,E., Lewis,L.C., Lewis,L., Li,J., Li,Z., Licharge,O., Lieu,C., Liu,J., Liu,W., Loulseged,H., Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J., Maheshwari,M., Mapa,P., Marondel,I., Martin,R., Martindale,A., Martinez,E., Massey,E., Maubley,E., McLeod,M.P., Meador,M., Mei,G., Merscher,S., Metzger,M., Miller,A., Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Montgomery,K.T., Morgan,M., Morris,S., Moser,M., Neal,D., Nelson,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N., Nguyen,N., Nickerson,E., Nwokkenko,S., Oguh,M., Okwumodu,G., Oragunde,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L., Peters,L., Pickens,R., Primus,E., Pu,L.L., Ruliz,S., Ren,Y., Rives,M., Rojas,A., Rojoubkan,I., Rolfe,M., Ruiz,S., Savery,G., Scherer,S., Scott,G., Shen,H., Shim,C., Shooshitari,N., Sisson,I., Sodergren,E., Sonalke,T., Sparks,A., Stanley,H., Stone,H., Sutton,A., Svatek,A., Taber,P., Tamerisa,A., Tamerisa,K., Tang,H., Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N., Thomas,S., Usmani,K., Vasquez,L., Vera,V., Villalon,D., Vanson,R., Wall,R., Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S., Williams,G., Williamson,A., Wleczyk,R., Wooden,K., Worley,K., Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorilla,S., Zucherlapati,R. and Gibbs,R.
 Direct Submission
 Unpublished
 2 (bases 1 to 176600)
 Worley,K.C.
 Direct Submission
 Submitted (13-APR-2000) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
 3 (bases 1 to 176600)
 Worley,K.C.
 Direct Submission
 Submitted (01-DEC-2000) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
 4 (bases 1 to 176600)
 Worley,K.C.
 Direct Submission
 Submitted (18-MAR-2003) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One

REFERENCE
 AUTHORS
 TITLE
 JOURNAL

COMMENT

Baylor Plaza, Houston, TX 77030, USA
On Dec 1, 2000 this sequence version replaced gi:11415072.
INFORMATION: <http://www.hgsc.bcm.tmc.edu/> or email
gc-help@bcm.tmc.edu

Phrap Value Range

CLONE LENGTH: This sequence does not necessarily represent the
entire length of this clone. Overlapping regions of clones are only
sequenced and submitted once, so the sequence for the remainder of
the insert may be found in the record for the adjacent clones.
Overlapping clones are noted at the beginning and end of the
features listing.

ANNOTATION OF FEATURES:

STS are identified using ePCR (Genome Res. 7:541-550) searches
of a local database that includes entries from dbSTS, GDB, and
local mapping efforts.

Repeats are identified using RepeatMasker (A. Smit and P. Green,
unpublished.) for Human and Mouse sequences.

Genes and Region of sequence similarity are identified by BLAST
(Nuc. Acids Res. 25:3389-3402) similarity (expect < 1e-34) to the
EST and cDNA sequences. Genes demonstrate at least two exons
flanked by consensus splice sites that maintained sequence
continuity across the splice junctions. Sequences that are not
identical matches are annotated as similar.

SEQUENCING READ COVERAGE: Sequencing is completed to a minimum
standard of double strand coverage with a minimum of 2 clones and 2
reads with no ambiguities or 2 chemistries with a minimum of 2
clones and 3 reads with no ambiguities. If the sequence quality for
a region does not meet this standard, it will be indicated in the
annotation as low coverage.

QUALITY OF INDIVIDUAL BASES: This sequence meets stringent quality
standards - estimated error rate less than 1 per 10,000 bases.
Reports of lowest quality individual bases and measures of base
quality are listed below. Description of the metrics can be found
at URL:
<http://gc.bcm.tmc.edu:8088/quality.info/genbank.annotation.html>.

QUALSTAT-REPORT-----

----- Summary Statistics -----
Contig length: 176600
Phrap values in estimate: 175849
Average error rate (BCM-Phrap estimate): 1.51359e-05
Fraction of Phrap values less than 40 : 0.0100029
Number of consensus changing edits: 3
Number of N's in consensus : 0

----- Consensus changing edits -----
Position Original+Context Edited+Context
48814 cctctgccta(n)ctgttagac cctctgccta(c)ctgttagac
48821 ctactctgta(n)gacttcacct ctactctgta(g)gacttcacct
158405 cctctctct(n)atataatata cctctctct(a)atataatata

----- Distribution of Quality < 40 Bases -----

10001	*						
9001	*						
8001	*						
7001	*						
6001	*						
5001	*						
4001	*						
3001	*						
2001	*						
1001	*						
01	*						
5	10	15	20	25	30	35	40

FEATURES
source

Version: 1.01 gxf.

repeat_region	location/Qualifiers
repeat_region	1..176600
repeat_region	/organism="Homo sapiens"
repeat_region	/mol_type="genomic DNA"
repeat_region	/db_xref="taxon:9606"
repeat_region	/chromosome="12"
repeat_region	/clone="RP11-847H18"
repeat_region	607..992
repeat_region	/rpl_family="MIT1A2"
repeat_region	1220..1627
repeat_region	/rpl_family="MSTD"
repeat_region	3841..4334
repeat_region	/rpl_family="MTLD"
repeat_region	5769..5973
repeat_region	/rpl_family="MTT13"
repeat_region	7338..7711
repeat_region	/rpl_family="MER52A"
repeat_region	7887..7909
repeat_region	/rpl_family="AT_rich"
repeat_region	8070..8090
repeat_region	/rpl_family="AT_rich"
repeat_region	8661..8881
repeat_region	/standard_name="G51086"
repeat_region	/db_xref="dbSTS:73714"
repeat_region	9908..10030
repeat_region	/rpl_family="MIR"
repeat_region	10613..10965
repeat_region	/rpl_family="THE1B"
repeat_region	10966..11237
repeat_region	/rpl_family="THE1B-internal"
repeat_region	complement(11240..13261)
repeat_region	/rpl_family="LIM2"
repeat_region	13358..13766
repeat_region	/rpl_family="MSTA"
repeat_region	14250..14387
repeat_region	/rpl_family="MIR"
repeat_region	14392..14692
repeat_region	/rpl_family="MER46C"
repeat_region	complement(14786..15147)
repeat_region	/rpl_family="THE1B"
repeat_region	complement(15635..15778)
repeat_region	/rpl_family="AluY"
repeat_region	complement(15941..16369)
repeat_region	/rpl_family="MSTB"
repeat_region	complement(17408..17727)
repeat_region	/rpl_family="L2"
repeat_region	18185..18226
repeat_region	/rpl_family="AT_rich"
repeat_region	18311..18837
repeat_region	/rpl_family="MER89"
repeat_region	19039..19547

Query Match 69.9%; Score 1298.4; DB 9; Length 176600;

Best Local Similarity 95.4%; Pred. No. 7.1e-269;

Matches 1325; Conservative 7; Mismatches 56; Indels 1; Gaps 1;

QY	366	CAAGCTGCTCCATCAGATATCCAGAGAGCTTGTCAACATTAACTAGCAGTATTTCACAC	425
DB	87687	CAAGCTGCTCCATCAGATATCCAGAGAGCTTGTCAACATTAACTAGCAGTATTTCACAC	87746
QY	426	ACTTATTTTCTTAACTCCCTTTTATTAACAGAGCAGATCTATGTTTTTACCTTTT	485
DB	87747	ACTTATTTTCTTAACTCCCTTTTATTAACAGAGCAGATCTATGTTTTTACCTTTT	87806
QY	486	TGCATATTTGATGTCTTTATGGAATCATAAATCTCAGCCTTCTTGATTTTGTGG	545
DB	87807	TGCATATTTGATGTCTTTATGGAATCATAAATCTCAGCCTTCTTGATTTTGTGG	87866

QY	546	CTTCATGTTGGCAAAACAATATCATCTGGGCTGCTCTCTGACGGGAATGTCATTGC	605
Db	87867	CTTCATGTTGGCAAAACAATATCATCTGGGCTGCTCTCTGACGGGAATGTCATTGC	87928
QY	606	ACAAAAGTATACAGAGCTTGGAAAACTGAGCTACAAAAGGAAGAGACAGACTCCACC	665
Db	87927	ACAAAAGTATACAGAGCTTGGAAAACTGAGCTACAAAAGGAAGAGACAGACTCCACC	87988
QY	666	TATTAAGGACCAATTGACAAATTCAGAAAAATTCAGATTCTTTGGCTTATTCAT	725
Db	87987	TATTAAGGACCAATTGACAAATTCAGAAAAATTCAGATTCTTTGGCTTATTCAT	88048
QY	726	GTCCTTTAAAACTGAGTATGCTTTTCGTGTGACTTGCGCCACATCCTTCGGGATT	785
Db	88047	GTCCTTTAAAACTGAGTATGCTTTTCGTGTGACTTGCGCCACATCCTTCGGGATT	88106
QY	786	TCTCTTTGAGCTTTTGTAGTATGATAGTGGAATGTTATGGGAGTGGAGTGA	845
Db	88107	TCTCTTTGAGCTTTTGTAGTATGATAGTGGAATGTTATGGGAGTGGAGTGA	88166
QY	846	TGAAGCCTGCTTCATTTTCCCTCAACTATTCATCTTTTTCATTTACTCTCTTTTTC	905
Db	88167	TGAAGCCTGCTTCATTTTCCCTCAACTATTCATCTTTTTCATTTACTCTCTTTTTC	88222
QY	906	TTTTTCCATCTCTGCTGCTCTTACGAAAATTAAGACTTTTCTTCTTACGTTGGAACA	965
Db	88227	TTTTTCCATCTCTGCTGCTCTTACGAAAATTAAGACTTTTCTTCTTACGTTGGAACG	88286
QY	966	TGGAATTCGTGTTTTGGTGGTTCACCTAGCTGCTGGTPTTATGGTTGGAAATTCACCTA	1025
Db	88287	TGGAATTCGTGTTTTGGTGGTTCACCTAGCTGCTGGTPTTATGGTTGGAAATTCACCTA	88344
QY	1026	TGCTCATTAATCTTGGCTAGACAGACATATAGACTTTACTTTCTATGTGTGAAAAAGAT	1085
Db	88347	TGCTCATTAATCTTGGCTAGACAGACATATAGACTTTACTTTCTATGTGTGAAAAAGAT	88406
QY	1086	TTTTCAAAAGTATAGCAATTCGAAATTTGTTGATGTTCCAGCCTATATATTCGTGGTG	1145
Db	88407	TTTTCAAAAGTATAGCAATTCGAAATTTGTTGATGTTCCAGCCTATATATTCGTGGTG	88466
QY	1146	GAGATATGCGACATCGAATTCGAAATTCGAAATTTTGGAAATTTATGTTTTGCATATG	1205
Db	88467	GAGATATGCGACATCGAATTCGAAATTCGAAATTTTGGAAATTTATGTTTTGCATATG	88522
QY	1206	CTTGCTATGTTATAGTTCCTCAGAACTGCGGGAATTCGTTACTTCATTTTACCTTA	1265
Db	88527	CTTGCTATGTTATAGTTCCTCAGAACTGCGGGAATTCGTTACTTCATTTTACCTTA	88588
QY	1266	TGTCATTTATATGCTTATACATTAACCTGCTCCACATCCAGACTGTTGTTGACTGAG	1325
Db	88587	TGTCATTTATATGCTTATACATTAACCTGCTCCACATCCAGACTGTTGTTGACTGAG	88644
QY	1326	TTGCTATGCAATGTTATTCATTAACCTTTTTCACATCTTCTGACAAAGCTTTGACAG	1385
Db	88647	TTGCTATGCAATGTTATTCATTAACCTTTTTCACATCTTCTGACAAAGCTTTGACAG	88706
QY	1386	GCCAAATATGTCAGACATTCAAAGGTTTATGTGTAATATACAGTATTTTGAACGTGA	1445
Db	88707	GCCAAATATGTCAGACATTCAAAGGTTTATGTGTAATATACAGTATTTTGAACGTG	88766
QY	1446	AAAAATGACCTTAATA-TAGACATTTCTACAAAGAACAGTAAATATGNGGAAAAATG	1504
Db	88767	AAAAATGACCTTAATAATAAGCACTTCTACAAAGAACAGTAAATATGNGGAAAAATG	88822
QY	1505	GAAATTCCTTTAGTGTGAGTGGGCTGCTCAAAATTCATATGTTTTTTTAAATATATT	1564
Db	88827	GAAATTCCTTTAGTGTGAGTGGGCTGCTCAAAATTCATATGTTTTTTTAAATATATT	88888
QY	1565	TAAACATATGTAAAGAAATTAAGTGGCAAAAGCTGGGAAAGCTTAAGACCTGCTCAAA	1624
Db	88887	TAAACATATGTAAAGAAATTAAGTGGCAAAAGCTGGGAAAGCTTAAGACCTGCTCAAA	88944
QY	1625	GCCGTAAATATGGGAAAAATAAANNNGTTTCACATATCTCATATCGCTCANNKATGNTG	1684

Db	Accession	Sequence	Length	Quality	Reference
Db	88947	GCCTGAAATATGCAAAATATTAATTTGTTTTCACATATATCATATCTCATATGTTG	1744	89006	
Qy	1685	GCCTGAAATATGCAAAATATTAATTTGTTTTCACATATATCATATCTCATATGTTG	1744	89006	
Db	89007	GCCTGAAATATGCAAAATATTAATTTGTTTTCACATATATCATATCTCATATGTTG	1744	89006	
Qy	1745	MMATNACT 1753			
Db	89067	CATATACT 89075			
RESULT 9	AC024478/c	158647 bp	DNA	linear	HTG 26-MAY-2000
DEFINITION	AC024478	Homo sapiens chromosome 12 clone RP11-776P23 map 12, WORKING DRAFT			
ACCESSION	AC024478	GI:8076945			
KEYWORDS	HTG; HTGS_PHASE1; HTGS_DRAFT.				
ORGANISM	Homo sapiens (human)				
REFERENCE	Anderson, S., Baldwin, J., Barra, N., Beda, F., Boguslavsky, L., Boukhalter, B., Brown, A., Burkett, G., Campomiano, A., Castle, A., Choepel, Y., Collangelo, M., Collins, S., Collymore, A., Cooke, P., DeRellano, K., Dewar, K., Diaz, J. S., Dodge, S., Domino, M., Doyle, M., Feneberg, J., Ferreira, P., Fitzhugh, W., Forrest, C., Gage, D., Galagan, J., Gardyna, S., Ginde, S., Goyette, M., Graham, L., Grand-Pierre, N., Grant, G., Hagos, B., Heaford, A., Horton, L., Howland, J. C., Iliev, I., Johnson, R., Jones, C., Kann, L., Karatas, A., Klein, J., Landers, T., Largocque, K., Lehoczy, J., Levine, R., Liu, G., Locke, K., MacDonald, P., Margis, N., McCarthy, M., McMan, P., McGurk, A., McKernan, K., McPheeters, R., Meldrum, J., Meneus, L., Mihova, T., Miranda, C., Mlenga, V., Morrow, J., Notman, C. H., O'Connor, T., O'Donnell, P., O'Neill, D., Olivar, T. M., Peterson, K., Pierre, N., Pisani, C., Pollara, V., Raymond, C., Riley, R., Rogov, P., Rothman, D., Roy, A., Santos, R., Schauer, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J., Testafay, S., Theodore, J., Tirrell, A., Travers, M., Trigilio, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G., Zainoun, J., Zimmer, A. and Zody, M.				
TITLE	Direct Submission				
JOURNAL	Submitted (28-FEB-2000) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA				
COMMENT	On May 25, 2000 this sequence version replaced gi:7259764. All repeats were identified using RepeatMasker: Smit, A.F.A. & Green, P. (1996-1997) http://ftp.genome.washington.edu/RM/RepeatMasker.html				
	Center: Whitehead Institute/ MIT Center for Genome Research				
	Center code: W1BR				
	Web site: http://www-seq.wi.mit.edu				
	Contact: sequence_submissions@genome.wi.mit.edu				
	Project Information				
	Center project name: L7124				
	Center clone name: 776.P.23				
	Summary Statistics				
	Sequencing Vector: M13; W77615; 100% of reads				
	Chemistry: Dye-terminator Big Dye; 100% of reads				
	Assembly program: Phrap; version 0.960731				
	Consensus quality: 135824 bases at least Q40				
	Consensus quality: 147226 bases at least Q30				
	Consensus quality: 152156 bases at least Q20				
	Insert size: 139000; agarose-1p				

Insert size: 155647; sum-of-contigs
Quality coverage: 3.6 in Q20 bases; agarose-fp
Quality coverage: 3.2 in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently
* consists of 31 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
1 1512: contig of 1512 bp in length
* 1513 1612: gap of 100 bp
* 1613 3017: contig of 1405 bp in length
* 3018 3117: gap of 100 bp
* 3118 4746: contig of 1629 bp in length
* 4747 4846: gap of 100 bp
* 4847 6511: contig of 1665 bp in length
* 6512 6611: gap of 100 bp
* 6612 8674: contig of 2063 bp in length
* 8675 8775 8774: gap of 100 bp
* 11190: contig of 2416 bp in length
* 11191 11290: gap of 100 bp
* 11291 14133: contig of 2843 bp in length
* 14134 14233: gap of 100 bp
* 14234 16333: contig of 2100 bp in length
* 16334 16433: gap of 100 bp
* 16434 20417: contig of 3984 bp in length
* 20418 20517: gap of 100 bp
* 20518 23097: contig of 2580 bp in length
* 23098 23197: gap of 100 bp
* 23198 27526: contig of 4329 bp in length
* 27527 27626: gap of 100 bp
* 27627 32176: contig of 4550 bp in length
* 32177 32276: gap of 100 bp
* 32277 36605: contig of 4329 bp in length
* 36606 36705: gap of 100 bp
* 36706 41220: contig of 4515 bp in length
* 41221 41321: gap of 100 bp
* 41321 44907: contig of 3587 bp in length
* 44908 45007: gap of 100 bp
* 45008 48806: contig of 3759 bp in length
* 48807 48906: gap of 100 bp
* 48907 53219: contig of 4313 bp in length
* 53220 53319: gap of 100 bp
* 53320 58369: contig of 5050 bp in length
* 58370 58469: gap of 100 bp
* 58470 62150: contig of 3681 bp in length
* 62151 62250: gap of 100 bp
* 62251 66960: contig of 4710 bp in length
* 66961 67060: gap of 100 bp
* 67061 72166: contig of 5106 bp in length
* 72167 72266: gap of 100 bp
* 72267 77323: contig of 5057 bp in length
* 77324 77423: gap of 100 bp
* 77424 82913: contig of 5490 bp in length
* 82914 83013: gap of 100 bp
* 83014 89483: contig of 6470 bp in length
* 89484 89583: gap of 100 bp
* 89584 97274: contig of 7691 bp in length
* 97275 97374: gap of 100 bp
* 97375 104650: contig of 7276 bp in length
* 104651 104751: gap of 100 bp
* 104751 112482: contig of 7732 bp in length
* 112483 112582: gap of 100 bp
* 112583 121328: contig of 8746 bp in length
* 121329 121428: gap of 100 bp
* 121429 128177: contig of 6749 bp in length
* 128178 128277: gap of 100 bp
* 128278 141975: contig of 13698 bp in length
* 141976 143075: gap of 100 bp
* 143076 158647: contig of 16572 bp in length.

FEATURES
source

misc_feature
1. .158647
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="12"
/map="12"
/clone="RP11-776p23"
/clone_11b="RP11-11 Human Male BAC"
1. .1512
/note="assembly_fragment"
misc_feature
1613. .3017
/note="assembly_fragment"
misc_feature
3118. .4746
/note="assembly_fragment"
misc_feature
4847. .6511
/note="assembly_fragment"
misc_feature
6612. .8674
/note="assembly_fragment"
misc_feature
8775. .11190
/note="assembly_fragment"
clone_end:77
vector_side:left
misc_feature
11291. .14133
/note="assembly_fragment"
misc_feature
14234. .16333
/note="assembly_fragment"
misc_feature
16434. .20417
/note="assembly_fragment"
misc_feature
20518. .23097
/note="assembly_fragment"
misc_feature
23198. .27526
/note="assembly_fragment"
misc_feature
27627. .32176
/note="assembly_fragment"
misc_feature
32277. .36605
/note="assembly_fragment"
misc_feature
36706. .41220
/note="assembly_fragment"
misc_feature
41321. .44907
/note="assembly_fragment"
misc_feature
45008. .48806
/note="assembly_fragment"
misc_feature
48907. .53219
/note="assembly_fragment"
misc_feature
53320. .58369
/note="assembly_fragment"
misc_feature
58470. .62150
/note="assembly_fragment"
misc_feature
62251. .66960
/note="assembly_fragment"
misc_feature
67061. .72166
/note="assembly_fragment"
misc_feature
72267. .77323
/note="assembly_fragment"
misc_feature
77424. .82913
/note="assembly_fragment"
misc_feature
83014. .89483
/note="assembly_fragment"
misc_feature
89584. .97274
/note="assembly_fragment"
misc_feature
97375. .104650
/note="assembly_fragment"
misc_feature
104751. .112482
/note="assembly_fragment"
clone_end:SP6
vector_side:left
misc_feature
112583. .121328
/note="assembly_fragment"

Query Match 66.4%; Score 1233.8; DB 2; Length 158647;
Best Local Similarity 88.9%; Pred. No. 5.7e-255;
Matches 1299; Conservative 3; Mismatches 159; Indels 1; Gaps 1;

QY 366 CAAGGCTGCCCTCAGATATCCAGAGAGCTCTTGCAACATTAACTAGCAGATATTTCCAC 425
DB 50198 CAAGGCTGCCCTCAGATATCCAGAGAGCTCTTGCAACATTAACTAGCAGATATTTCCAC 50139
QY 426 ACTTTATTTTAACTTCTTATATATCAGAGAGAGATCTATGTTTTTACTCTTTT 485
DB 50138 ACTTTATTTTAACTTCTTATATATCAGAGAGAGATCTATGTTTTTACTCTTTT 50079
QY 486 TGCATTTTGTGCTTTTATGGAATATATAAATCTGAGCTTCTGGAATTTTGTG 545
DB 50078 TGCATTTTGTGCTTTTATGGAATATATAAATCTGAGCTTCTGGAATTTTGTG 50019
QY 546 CTTTATGTTTGGCAACAATATCATCTGGGCTCTTCTGTGCAAGGATGTCATTC 605
DB 50018 CTTTATGTTTGGCAACAATATCATCTGGGCTCTTCTGTGCAAGGATGTCATTC 49959
QY 606 ACAAAGTTACTGAGGCTTGGAAGCTGACCTACAAAGAGAGACAGACTTCACC 665
DB 49958 ACAAAGTTACTGAGGCTTGGAAGCTGACCTACAAAGAGAGACAGACTTCACC 49899
QY 666 TATTAAGGACATTGGCAATTCGAAATTCCTCAGTTCTTGGCTTATTCAT 725
DB 49898 TATTAAGGACATTGGCAATTCGAAATTCCTCAGTTCTTGGCTTATTCAT 49839
QY 726 GTCTTTTAAACCTGAGTATGCTTTCTGTTGACTTGCCCTACATCTCTGGAT 785
DB 49838 GTCTTTTAAACCTGAGTATGCTTTCTGTTGACTTGCCCTACATCTCTGGAT 49779
QY 786 TCTGTTTGTGCTTTTGTAGTATATGTTGAATTTGTTATGGCATGAGTAGTCA 845
DB 49778 TCTGTTTGTGCTTTTGTAGTATATGTTGAATTTGTTATGGCATGAGTAGTCA 49719
QY 846 TGAAGCTGCTTCTTCTTCCCAACTATTTTCTTCTTCTTCTTCTTCTTCTTCTTCT 905
DB 49718 TGAAGCTGCTTCTTCTTCCCAACTATTTTCTTCTTCTTCTTCTTCTTCTTCTTCT 49659
QY 906 TTTTCTCATCTCTGCTCTCTAGCAAAATTAAGACTTCTTCTTCTTCTTCTTCTTCT 965
DB 49658 TTTTCTCATCTCTGCTCTCTAGCAAAATTAAGACTTCTTCTTCTTCTTCTTCTTCT 49599
QY 966 TGAATTCCTGTTTGTGCTTACCTTCTCTGCTTGTGTTTATGTTGAATTCCTTA 1025
DB 49598 TGAATTCCTGTTTGTGCTTACCTTCTCTGCTTGTGTTTATGTTGAATTCCTTA 49539
QY 1026 TGCCTAATAATCTTCTAGCAGACAAATGACATTTACTTCTTCTTCTTCTTCTTCT 1085
DB 49538 TGCCTAATAATCTTCTAGCAGACAAATGACATTTACTTCTTCTTCTTCTTCTTCT 49479
QY 1086 TTTTCAAGATATGCAATTCGAAATTTGTTAGTCCAGCCTATATTTGCTGCTG 1145
DB 49478 TTTTCAAGATATGCAATTCGAAATTTGTTAGTCCAGCCTATATTTGCTGCTG 49419
QY 1146 GAGTATAGCTGACTATTAATCAAGCCAAATTTTGTGAATTTATGTTTTTCAT 1205
DB 49418 GAGTATAGCTGACTATTAATCAAGCCAAATTTTGTGAATTTATGTTTTTCAT 49359
QY 1206 CTTGTTCAATTTATATGTTCTCCAGAAAGCTGGAATTTTCTTCTTCTTCTTCTTCT 1265
DB 49358 CTTGTTCAATTTATATGTTCTCCAGAAAGCTGGAATTTTCTTCTTCTTCTTCTTCT 49299
QY 1266 TGTCAATTTATAGCTTACATTAATCTGCTCCACATCCAGACTGTTTGTGAATG 1325
DB 49298 TGTCAATTTATAGCTTACATTAATCTGCTCCACATCCAGACTGTTTGTGAATG 49239
QY 1326 TGTCTATGCAATTTATTTATTAATCAATTTTATCATCTTCTGCAACAGACTTTG 1385
DB 49238 TGTCTATGCAATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 49179
QY 1386 GCGAATATGCTGAGCATTCGAAGGTTTATGTTATTTATGTTATTTATTTATTTAT 1445
DB 49178 GCGAATATGCTGAGCATTCGAAGGTTTATGTTATTTATGTTATTTATTTATTTAT 49119
QY 1446 AAAATGACTTATATA-TAGACCATTCTTACAAAGAACATGAATAGGNGAAAAATG 1504

DB 4918 AAAATGACTTATATAATAGACATTCTTACAAAGAACATGAATAGGNGAAAAATG 49059
QY 1505 GAATTTCTTTAGTGCAGAGGCTGCTTCAATTTACATTTAGTTTTTATATATATTT 1564
DB 49058 GAATTTCTTTAGTGCAGAGGCTGCTTCAATTTACATTTAGTTTTTATATATATTT 48999
QY 1565 TAAACATATGTAAGTAATTAAGTGGCAAGAACCTGGAAGCTTAAGACCTGCTCAAN 1624
DB 48998 TAAACATATGTAAGTAATTAAGTGGCAAGAACCTGGAAGCTTAAGACCTGCTCAAN 48939
QY 1625 GCGTGAATATGGAATAATAANNGTTTNCAGATATCTCATTCCTGCTNNNNATGNTG 1684
DB 48938 GCGTGAATAATGGAATAATAATTTGTTTTCANNNNNNNNNNNNNNNNNNNNNNN 48879
QY 1685 GCGCTTMCANAGCTTGGGAGTGTNNNNMTGATATAGTTNATTAANCTGGGNTGCTN 1744
DB 48878 NNN 48819
QY 1745 NNNATGCTTNNKNCANCMNNNNACATGNNNTANNANNTATTACAAAGCTCAGT 1804
DB 48818 NNNNNNNNNNNNTTACCATCAGTATATTTGTTGCTTATTCATCTTCCCTG 48759
QY 1805 GATATTCCTGACTGGAAGTGC 1826
DB 48758 TAGAATATTAAGTATAAAGGC 48737
RESULT 10
BD160531.
LOCUS BD160531 2362 bp DNA linear PAT 17-JAN-2003
DEFINITION Primer for synthesizing full-length cDNA and use thereof.
ACCESSION BD160531
VERSION BD160531.1 GI:27866289
KEYWORDS JP 2002191363-A/15374.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 2362)
Ota,T., Isogai,T., Nishikawa,T., Hayashi,K., Saito,K., Yamamoto,J.,
Ishii,S., Sugiyama,T., Wakamatsu,A., Nagai,K., and Otsuki,T.
Primer for synthesizing full-length cDNA and use thereof
Patent: JP 2002191363-A 15374 09-JUL-2002;
JOURNAL HELIX RESEARCH INSTITUTE
COMMENT OS Homo sapiens (human)
PN JP 2002191363-A/15374
PD 09-JUL-2002
PE 28-JUL-2000 JP 2000280990
PI TOSHIO OTA, TAKAO ISOGAI, TETSUO NISHIKAWA, KOJI HAYASHI, KAZUO
PI SATTO,
PI JUNICHI YAMAMOTO, SHIZUKO ISHII, TOMOYASU SUGIYAMA, AI WAKAMATSU,
PI KEIICHI NAGAI, TETSUO OTSUKI
PC C12N15/09, C07K14/47, C07K16/18, C12N1/15, C12N1/19, C12N1/21, C12N5/ PC
10, PC C12P21/02, C12Q1/68, C12P21/08, G06F17/30, C12N15/00, C12N5/00 CC
Primer for synthesizing full-length cDNA and use thereof FH key
FEATURES
source FT CDS Location/Qualifiers
1..2362
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
BASE COUNT 687 a 398 c 413 g 864 t
ORIGIN
Query Match 65.5%; Score 1217.2; DB 6; Length 2362;
Best Local Similarity 95.0%; Pred. No. 2.2e-251;
Matches 1245; Conservative 7; Mismatches 58; Indels 1; Gaps 1;
QY 444 CCTTATATATACGAAGCAGATCTATGTTTTTACTCTTTTGCATATTTGATGTGCT 503

[illegible]

	BASE COUNT	ORIGIN
Db	1081 GTGGTGGCTTCAAATTACATTAAGTTTTTTTTATATATATTTTAAACAATGTGAAGAAT	1149
Qy	1583 TAAGTGCCAAAGACTGGGAAAAGCTTAAAGACCCTGTTCAANCGCTGAATAATGGAAAA	1642
Db	1141 TAAGTGCCAAAGAACTGGAAAAGCTTAAAGACCCTGTTCAANCGCTGAATAATGGAAAAA	1200
Qy	1643 TAAAMNCTTTNCATATATCATATTCGCCTCANNKNAANGNGGCCCTTMNCAANCCTTGG	1702
Db	1201 TAAAAATTTGTTTCAATATATCATATTCACACTGCATAAAGATGGGCCCTTAAAAAACCTTGG	1260
Qy	1703 GAATGKTNNMNTGNATTAAGTTNATTAANCAANGGNGGNTGCTNNMATNACT	1753
Db	1261 GAATGTTTGTATGTACAAAGTTATTTAAACGCGGTATGCTTCATATTACT	1311
RESULT 11 AK027657 LOCUS	AK027657	2362 bp mRNA linear PRI 01-AUG-2002
DEFINITION	Homo sapiens CDNA FLJ14751 fis, clone NT2RP3003059, moderately similar to Rattus norvegicus potassium channel regulator I mRNA.	
ACCESSION	AK027657	
VERSION	AK027657.1	GI:14042497
SOURCE	oligo capping; fis (full insert sequence).	
ORGANISM	Homo sapiens (human)	
REFERENCE AUTHORS	Homo sapiens Eutheria; Metazoa; Chordata; Craniata; Vertebrate; Euteleostomi; Mammalia; Euteria; Primates; Carnivora; Hominiidae; Homo. 1 Isogai,T., Ota,T., Hayashi,K., Sugiyama,T., Otsuki,T., Suzuki,Y., Nishikawa,T., Nagai,K., Sugano,S., Takahashi-Fujii,A., Hara,H., Tanabe,T., Nomura,Y., Togiyasu,S., Komai,F., Hara,R., Takeuchi,K., Arita,M., Nabekura,T., Ishii,S., Kawai,Y., Saito,K., Yamamoto,J., Wakamatsu,A., Nakamura,Y., Nagahari,K., Masubo,Y. and Oshima,A. NEBO human cDNA sequencing project unpublished 2 (bases 1 to 2362) Isogai,T. and Otsuki,T. Direct Submission Submitted (10-MAY-2001) Takao Isogai, Helix Research Institute, Genomics Laboratory, 1532-3 Yana, Kisarazu, Chiba 293-0812, Japan (E-mail:genomics@helix.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986) NEBO human cDNA sequencing project supported by Ministry of Economy, Trade and Industry of Japan. CDNA full insert sequencing: Research Association for Biotechnology: cDNA library construction, 5'- & 3'-end one pass sequencing and clone selection: Helix Research Institute (supported by Japan Key Technology Center etc.) and Department of Virology, Institute of Medical Science, University of Tokyo. Location/Qualifiers 1. 2362 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="NT2RP3003059" /cell_line="NT2" /cell_type="teratocarcinoma" /note="cloning vector: pME18SFL3-mRNA from NT2 neuronal precursor cells after 2-weeks retinoic acid (Ra) induction." 16 . 979 /note="unnamed protein product" /codon_start=1 /protein_id="BA855272.1" /db_xref="GI:14042498" /translation="MFPTLFAVYLICLYGNHKTSALFGCGFMPEROTNIWAVOCAGNVIAOITLPAWKTELOKKEDRLPIYGCPAEFRKIIOPLIAYSMSRKNLSMLTLTPYIILGLIFCAPVINGVIIGIDRSSHEACILHPLEFSPILFESEPHLLPSPKIKFPLSIWKRRILFEVVIVLVSVLPVWKRTVAHKYLLADNRNRYTVMKRVOREETKYLIYPAYFAEMASTDAISKSIFWNLMFFCIETTYIVIPOLUEPRFIILEPYIRKLNPDPPTSILCELSTDAIVANVTITTFILNKFTQMPNSODIQREWM"	

TITLE	JOURNAL	REFERENCE	AUTHORS
Morgan, M., Morris, K., Morris, S., Mundias, M., Murphy, M., Nair, L., Markert, C.C., Neal, D., Newton, N., Nguyen, N., Norris, S., Parks, K., Mwokeremah, O., Okunodu, G., Olarinnasogoon, A., Pal, S., Patks, K., Paternak, S., Paul, H., Perez, A., Perez, I., Pfannkuch, C., Plapper, F., Poindexter, A., Popovic, D., Piums, E., Pu, L., Puzo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M.A., Reihl, R., Kelly, B., Kelly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F., Rivers, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S.J., Sanders, W., Savery, G., Scherer, S., Scott, G., Shatsman, S., Shen, H., Shetty, J., Shvartsbeyn, A., Siesson, I., Sliter, C.D., Smayls, D., Sneed, A., Sodergren, E., Song, X.-Z., Stokell, R., Sosa, J., Steelma, M., Strong, R., Sutton, A., Svatek, A., Taber, P., Taylor, T., Thomas, N., Thoms, S., Tingey, A., Trejos, Z., Usmani, K., Valas, R., Veta, V., Villasana, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F., Williams, G., Willson, R., Wleczek, R., Wooden, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, Y., Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von Niederhausen, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O., Weinstock, G., and Gibbs, R.A.	Unpublished	2 (bases 1 to 276372)	Worley, K.C.
Direct Submission	Submitted (29-JUN-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA	3 (bases 1 to 276372)	Rat Genome Sequencing Consortium.
Direct Submission	Submitted (03-OCT-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA	On Sep 23, 2002 this sequence version replaced gi:21628982.	The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (http://www.hgsc.bcm.tmc.edu/projects/rat/). As a result, the sequence may extend beyond the ends of the clone and there may be contigs that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.
Genome Center:	Baylor College of Medicine	Center code: BCM	Web site: http://www.hgsc.bcm.tmc.edu/
Contact:	hgsc-help@bcm.tmc.edu	Project Information	Center project name: GDBR
Center clone name:	CH230-8B24	Summary Statistics	Assembly program: Phrap; version 0.990329
Consensus quality:	193368 bases at least Q40	Consensus quality:	199764 bases at least Q30
Consensus quality:	203624 bases at least Q20	Estimated insert size:	222869; sum-of-contigs estimation
Quality coverage:	4x in Q20 bases; sum-of-contigs estimation	Note:	NOTE: Estimated insert size may differ from sequence length
(see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html)	* NOTE: This is a 'working draft' sequence. It currently consists of 9 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown.	* This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.	* 1 4680: contig of 4680 bp in length
* 4681 4780: gap of unknown length	* 4781 243181: contig of 238401 bp in length	* 243182 247497: gap of unknown length	* 243182 247497: contig of 4216 bp in length

```

Db      140447  ACGTAGAGTCAGTCTCTGTGANTACTTAACTGCTGGTATTTTGGTTGCAAAATCAC 140506
Oy      1023  TTATGTCATATAAATCTTCTGTAAGACAGACATAGACATTAATCTTCTATGTGTGGAAG 1082
Db      140507  TTAGTCATATAGATTTTACGTACGACAGACATAGACATTAATGATTTTATGTGTGGAAG 140566
Oy      1083  AGTTTTCAGATATAGCAATTTCTGGAATATTTTGTAGTCTCCAGCCTATATATTTGCTGG 1142
Db      140567  AGTATTTTCAGACACATGAATATGTCATATTTATTTAGTCCAGCCTCATGTTTGTGCG 140626
Oy      1143  TTGGAGTATAGTCATGCTATGAATATCAAGCCATTTTGGATTAATTAAGTTTTCAT 1202
Db      140627  TTGGGCTTACCTGACCTCTTAAATTCAGAAAGCAATTTTCGGAATTTAAAGTTTGT 140686
Oy      1203  ATGCTGTTCATTTGTTATGATTCCTCCAGAAACTGCTGGAATTTCTACTCATTTTACC 1262
Db      140687  ATGCTGTGTTGCTTCTTACGGTTCCTCCAGAAACTACTAGATTTCCGTTACTTCAATTTAC 140746
Oy      1263  TTATGTCATTTATAGGCTTAAACATACCTGCTGCCACATCCAGACATTTGTTGAAC 1322
Db      140747  ATACATTTATTTATAGCTCAGACATGCTCTGCGACCCATATCTAGACTTGTGTGAAC 140806
Oy      1323  GAGTTGCTATGCAATGTTATTTATTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 1382
Db      140807  GGGTTGCTATGCAATTTGTTCAATTTTCTACATTTTCTTGAACCAAGACTTTTCA 140866
Oy      1383  GTGGCCAATATGTCAGACATTTCAAGGTTTATGTGAATATTCAGTATATTTTGAAC 1442
Db      140867  GTGGTCGATATGTCATGACATCCAAAGTTTATGTGTAGTGTGAGGCTATTTTGAAC 140926
Oy      1443  GTAAATATGACTT 1456
Db      140927  CTAAGAGCCTATTT 140940

RESULT 14
AC114560/c 152534 bp DNA linear HTG 11-JUN-2003
LOCUS Mus musculus clone RP23-54K6, *** SEQUENCING IN PROGRESS ***, 3
ORDERED pieces.
AC114560
AC114560.5 GI:31581666
HTG: HTGS_PHASE2; HTGS_FULFILLTOP; HTGS_ACTIVEFIN.
KEYWORDS Mus musculus (house mouse)
SOURCE Mus musculus
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 152534)
AUTHORS Birren, B., Nusbaum, C. and Lander, E.
TITLE Mus musculus, clone RP23-54K6
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 152534)
AUTHORS Birren, B., Linton, L., Nusbaum, C., Lander, E., Allen, N.,
Anderson, S., Barina, N., Bastien, V., Bloom, T., Boguslavsky, L.,
Bouhagalter, B., Brown, A., Camarata, J., Campiano, A., Chang, J.,
Chazaro, B., Choepel, Y., Colangelo, M., Collins, S., Collymore, A.,
Cook, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J.S., Dodge, S.,
Faro, S., Ferreira, P., Fitzhugh, M., Gage, D., Galagan, J., Gardyna, S.,
Gilde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N.,
Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C.,
Kamat, A., Karatas, A., Kells, C., Lacroque, K., Lamazares, R.,
Lander, T., Lehoczy, J., Levine, R., Lindblad-Toh, K., Liu, G.,
Maclean, C., Macdonald, P., Major, J., Marquis, N., Matthews, C.,
McCarthy, M., McEwan, P., McKernan, K., Meldrum, J., Meneus, L.,
Mihova, T., Mienga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R.,
Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D.,
O'Neil, J., Peterson, K., Phunhhang, P., Pierre, N., Pollara, V.,
Raymond, C., Rella, R., Rieback, M., Riley, R., Rise, C., Rogov, P.,
Roman, J., Roselli, M., Roy, A., Santos, R., Schauer, S., Schuback, R.,
Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N.,
Strauss, K., Subramanian, A., Talamas, J., Testfaye, S., Theodore, J.,
Topham, K., Travers, M., Trivis, N., Trigglio, J., Vassiliev, H.,
Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, G.,

```

```

TITLE Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.
JOURNAL Direct Submission
REFERENCE Submitted (10-Mar-2002) Whitehead Institute/MIT Center for Genome
AUTHORS Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 152534)
Birren, B., Nusbaum, C., Lander, E., Abouelkell, A., Allen, N.,
Anderson, S., Arachchi, H.M., Barina, N., Bastien, V., Bloom, T.,
Boguslavsky, L., Bouhagalter, B., Camarata, J., Chang, J., Choepel, Y.,
Collymore, A., Cook, A., Cooke, P., Corum, B., Dearellano, K.,
Diaz, J.S., Dodge, S., Dooley, K., Dorris, L., Erickson, J., Faro, S.,
Ferreira, P., FitzGerald, M., Gage, D., Galagan, J., Gardyna, S.,
Graham, L., Grand-Pierre, N., Hafez, N., Hagopian, D., Hagos, B.,
Hall, J., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C.,
Kamat, A., Karatas, A., Kells, C., Lander, T., Levine, R.,
Lindblad-Toh, K., Liu, G., Lui, A., Mabbitt, R., Maclean, C.,
Macdonald, P., Major, J., Manning, J., Matthews, C., McCarthy, M.,
Meldrum, J., Meneus, L., Mihova, T., Mienga, V., Murphy, T., Naylor, J.,
Nguyen, C., Nicol, R., Norbu, C., O'Connor, T., O'Donnell, P.,
O'Neill, D., Oliver, J., Peterson, K., Phunhhang, P., Pierre, N.,
Rachupka, A., Ramasamy, U., Raymond, C., Rella, R., Rise, C., Rogov, P.,
Roman, J., Schauer, S., Schuback, R., Seaman, S., Severy, P., Smith, C.,
Spencer, B., Stange-Thomann, N., Stojanovic, N., Stubbs, M.,
Talamas, J., Testfaye, S., Theodore, J., Topham, K., Travers, M.,
Vassiliev, H., Venkataraman, V.S., Viel, R., Vo, A., Wilson, B., Wu, X.,
Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.
Direct Submission
Submitted (11-Jun-2003) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Jun 11, 2003 this sequence version replaced gi:30984906.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: MIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L23271
Center clone name: 54_K_6
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 3 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* provided by the submitter.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
* 1 80469: contig of 80469 bp in length
* 80470 80569: gap of 100 bp
* 80570 90010: contig of 9441 bp in length
* 90011 90110: gap of 100 bp
* 90111 152534: contig of 62424 bp in length.
Location/Qualifiers
1. 152534
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/clone="RP23-54K6"
/clone_11b="RP21-23 Female Mouse BAC"
BASE COUNT 44923 a 30812 c 30537 g 45587 t 675 others
ORIGIN
Query Match 40.2% Score 746.8 DB 2: Length 152534;
Best Local Similarity 81.1% Pred. No. 3.4e-150;
Matches 881; Conservative 0; Mismatches 202; Indels 3; Gaps 1;
Oy 366 CAAGCTGCTTCATATGACAGAGAGTCTTGTACACATTAACACAGAGATTTTCCAC 425
Db 52439 CAAGCTTCTTCATATGACAGAGATCTTGTACACATTAACCTTACAGTCTTCCAC 52380

```

QY	426	ACTTATTTTAACTTCCTTATATATACAGACGAGATCTAGTCTTTTACTCTTTT	485
Db	52379	CCCTATTTTTTTTAACTTCCTTACATATACAGAGCTGGCTGTATTTCTTACTCTTTT	523200
QY	486	TGCATATATATATGTCCTTATATGGAATCATATAACTTCAGCCTTCCTTGAAATTTTGCG	545
Db	52319	TGCTTATTTGATAGTGTCTTATATGCAACCATAGACCTGGCCTTGCTTGGAATTTTGCG	52260
QY	546	CTTCATCTTTTGGCAACAATATCATCTGGCTGTCTCTGTGACGGAAATGCAATTC	605
Db	52259	TTTTCATCTTTTGGCAACAATATCATCTGGCTGTCTCTGTGACGGAAATGCAATTC	522000
QY	606	ACAAAGCTTAACTGAGGCTTGGAAAAGTGAAGTCAAAA-----AAGAGAGAGACAGACTTC	662
Db	52139	ACAGAACTGCAAGTGAAGCTTGGAAATTTGAGTGTGCGAGAAAGAAAGAGAGAGCTTC	521400
QY	663	ACCTATTTAAAGACCAATTTGCAGAAATTCAGAAAAATTTCTAGTTCTTTTGGCTTATTC	722
Db	52139	GCCCCAAGAGGAGCCGCTCGGAGCTCAGAAAGAGTGTGACAGTTTCTCTGTGTGACGC	520800
QY	723	CATGCTCTTAAAAACTTGAGTATAGCTTTTCTGTGTGACTTGGCCCTACATCCTTCTGG	782
Db	52079	CATGCTCTTAAAAAGCACTGAGAAATGCTCTTCTGTGACCTGGCCCTACGTCCTTCTGT	520200
QY	783	ATTCTCTTTTGTGCTTTTGTAGTATGTAATGCTGGAATTTGTAATGGAGATGAGTATG	842
Db	52019	GTTGGCGTCTTTTGGCTTTTGTGATGTATAGTGTGATGCTGTTGTGGTGGACCGAGTATG	519600
QY	843	TCATGAAGCCTGTCTTCATTTTCTCTCAACTATTTCACTTTTTCATTTACTCTCTTTT	902
Db	51959	TCAGAGAGCCGTCTCTCATTTTCTCAGTTGTTCTTTCTCTCTCTTACGACCTTTT	519000
QY	903	TTCTTTTCCATCT	962
Db	51899	CTCTTCT	518400
QY	963	ACATGGAATTCGTGTTTTGGTGTACCTTATGCTGTGTGTTTTAGTTGGAAATTCAC	1022
Db	51839	GCGTAGAGCTCAGATTCCTGTGTGTACGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	517800
QY	1023	TTATGCTCATTAATATCTTGTGTACAGAAATATACAAATTAATCTTCTATGTGTGAAAAG	1082
Db	51779	TTATGCTCATTAATATTAATCTGTGTACAGAAATATACAAATTAATCTTCTATGTGTGAAAAG	517200
QY	1083	AGTTTTTCAAGAGATATGCAATTTGTGAATTTTGTATGTATCCAGCCTATATATTTGCTGG	1142
Db	51719	AGATTTTCAAGAGATATGCAATTTGTGAATTTTGTATGTATGTATCCAGCCTATATATTTGCTGG	516600
QY	1143	TTGGAGTATAGCTGACTCATTTGAATTCAAAGCCAAATTTTTGGAAATTAATGTTTTTCAT	1202
Db	51659	TTGGAGTATAGCTGACTCATTTGAATTCAAAGCCAAATTTTTGGAAATTAATGTTTTTCAT	516000
QY	1203	ATGCTTCTTCATATGTTTATGTTTCTCTAGAAAGCTGGGAATTCGTTACTCTATTTTTCAC	1262
Db	51599	ATGCTTCTTCATATGTTTATGTTTCTCTAGAAAGCTGGGAATTCGTTACTCTATTTTTCAC	515400
QY	1263	TTATGTCATTTAAGGCTTACATACTGTGCTCCACATCCAGACTGTGTTGTGAAT	1322
Db	51539	ATACATATTTAATTAAGGCTTACATACTGTGCTCCACATATCTAGACTGTGTTGTGAAT	514800
QY	1323	GAGTTGCTATGCAATTTCTTATTTTCATACTTTTACATCTTTTCTGAAACAAAGCTTTTCA	1382
Db	51479	GAGTTGCTATGCAATTTCTTATTTTCATACTTTTACATCTTTTCTGAAACAAAGCTTTTCA	514200
QY	1383	GTGGCCAAATATGTCAGACATTCGAAAGTTTATGTGTAATATACATGATATTTTGAAGT	1442
Db	51419	GTGGCCAAATATATGTCAGACATTCGAAAGTTTATGTGTAATATATATTTTGAAGT	513600
QY	1443	GTAAAA 1448	
Db	51359	TTTAAA 51354	

LOCUS	BD148636	807 bp	DNA	linear	PAT 17-JAN-2003
DEFINITION	Primer for synthesizing full-length cDNA and use thereof.				
ACCESSION	BD148636	1	GI:27854394		
KEYWORDS	JP 2002191363-A/3479.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.				
AUTHORS	Ota,T., Isogai,T., Nishikawa,T., Hayashi,K., Saito,K., Yamamoto,J., Ishii,S., Sugiyama,T., Wakamatsu,A., Nagai,K. and Otsuki,T.				
TITLE	Primer for synthesizing full-length cDNA and use thereof				
JOURNAL	Patent: JP 2002191363-A 3479 09-JUL-2002;				
COMMENT	HELIIX RESEARCH INSTITUTE OS Homo sapiens (human) JP 2002191363-A/3479 PD 09-JUL-2002 PF 28-JUL-2000 JP 2000280990 PI TOSHIO OTA,TAKAO ISOGAI,TETSUO NISHIKAWA,KOJI HAYASHI,KAORU PI SAITO, PI JUNICHI YAMAMOTO,SHIZUKO ISHII,TOMOYASU SUGIYAMA,AI WAKAMATSU, PI KEIICHI NAGAI,TETSUJI OTSUKI PC C12N15/09,C07K14/47,C07K16/18,C12N1/15,C12N1/19,C12N1/21,C12N5/PC 10, PC C12P21/02,C12Q1/68/C12P21/08,G06F17/30,C12N15/00,C12N5/00 CC Primer for synthesizing full-length cDNA and use thereof FH Key Location/Qualifiers FT source 1..807 FT Location/Qualifiers 1..807 Location/Qualifiers /organism="Homo sapiens" /mol_type="genomic DNA" /db_xref="taxon:9606"				
FEATURES	source				
BASE COUNT	200 a 140 c 151 g 312 t 4 others				
ORIGIN					
Query Match	39.1%; Score 725.2; DB 6; Length 807;				
Best Local Similarity	95.4%; Pred. No. 1.6e-145;				
Matches	768; Conservative 0; Mismatches 34; Indels 3; Gaps 2;				
QY	444 CTTTATTATACAGACGAGATCTATGTTTCTTACTCTTTTGCATATTTGATGTCT 503				
DB	1 CCTTATTATACAGACGAGATCTATGTTTCTTACTCTTTTGCATATTTGATGTCT 60				
QY	504 TTATGGAAATCATAAACTTCAGCCCTTCCTTGGAATTTGTGGCTTCATGTTTGGCAAC 563				
DB	61 TTATGGAAATCATAAACTTCAGCCCTTCCTTGGAATTTGTGGCTTCATGTTTGGCAAC 120				
QY	564 AAATATCATCTGGCGCTCTCTCTGTGCAGGGAATGTCATTGCACAAAAAGTTAACTGAGC 623				
DB	121 AAATATCATCTGGCGCTCTCTCTGTGCAGGGAATGTCATTGCACAAAAAGTTAACTGAGC 180				
QY	624 TTGGAACACAGCTACACAAAGAAAGAACAGACACTTCACCTATTAAGAGCATTGGC 683				
DB	181 TTGGAACACAGCTACACAAAGAAAGAACAGACACTTCACCTATTAAGAGCATTGGC 240				
QY	684 AGAATTCGAAAAAATCTTCAGTTCTTTGGCTTATTCACATGTCCTTTAAAACTTGAG 743				
DB	241 AGAATTCGAAAAAATCTTCAGTTCTTTGGCTTATTCACATGTCCTTTAAAACTTGAG 300				
QY	744 TATGCTTTTCGTGTGACTTGGCCCTACATCCTTCCTGGAGTTTCTGTGTTGCTTTTGT 803				
DB	301 TATGCTTTTCGTGTGACTTGGCCCTACATCCTTCCTGGAGTTTCTGTGTTGCTTTTGT 360				
QY	804 AGTAGTTATAGTGTGAATTTGTTATTTGGCGATCGAGTAGATGATGAAGCCGTCTTCATTT 863				
DB	361 AGTAGTTATAGTGTGAATTTGTTATTTGGCGATCGAGTAGATGATGAAGCCGTCTTCATTT 420				

```
QY      864 TCCGACATATCTACTTTTTCATTTACTCTCTTTTCTTTCTTCTCTCACTCTCTGC 923
      |||
Db      421 TCCCAACTATCTACTTTTTCATTTACTCTCTTTTCTTTCTTCTCTCACTCTCTGC 480
QY      924 TCCAGCAAAATTTAGACTTTCTTCTCTAGTTTGAACATGSAATTCGTTTGT 983
      |||
Db      481 TCCTAGCAAAATTTAGACTTTCTTCTCTAGTTTGAACATGSAATTCGTTTGT 540
QY      984 GGTACCTTACTCTCTGTGTTTGAATTTGAAATTCATTAAGCTCAATAAATCTTGT 1043
      |||
Db      541 GGTACCTTACTCTCTGTGTTTGAATTTGAAATTCATTAAGCTCAATAAATCTTGT 600
QY      1044 AGCAGACATATGACATTTACTTCTATGTGTGAAAGAGTTTTCAAAGATATGCAAT 1103
      |||
Db      601 AGCAGACATATGACATTTACTTCTATGTGTGAAAGAGTTTTCAAAGATATGAAAC 660
QY      1104 TCTGAAATATTTGTAGTTCCAGCCTATATATTTGCTGTGAGATATAGCTCAAT 1163
      |||
Db      661 TGT-AAATATTTGTAGTTCCAGCCTATATATTTGCTGTGAGATATAGCTCAAT 719
QY      1164 GAAATCAAGCCAAATTTTGTGAATTTAATGTTTTCATATGCTT--GTTCATTGTATA 1221
      |||
Db      720 GAAATCAAGCCAAATTTTGTGAATTTAATGTTTTCATATGCTTGGGTCACTGTATAG 779
QY      1222 GTTCTCTGAAACTGCTGGAATTC 1246
      |||
Db      780 GTTCTCTGAAACTGCTGGAATTC 804
```

Search completed: September 23, 2003, 17:35:36
Job time : 22792 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 23, 2003, 11:15:44 ; Search time 44 Seconds
(without alignments)
2533.171 Million cell updates/sec

Title: US-10-000-151b-3
Perfect score: 6079
Sequence: 1 MPVRGHAAPONTFLDTIIR.....GOLGALTSQPLRHGSDPGS 1159

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	6079	100.0	1159	2	I38465
2	1767	29.1	514	2	T19579
3	1614	26.6	1017	2	T31354
4	1531	25.5	1102	2	T17367
5	1537.5	25.3	1087	2	T31100
6	1485.5	24.4	1284	2	T13168
7	1448.5	23.8	1174	2	A40853
8	1409.5	23.2	962	2	I53197
9	1402	23.1	989	2	I46912
10	1401	23.0	934	2	T42394
11	451.5	7.4	665	2	A552072
12	440	7.2	706	2	A55251
13	433	7.1	695	2	S74179
14	428.5	7.0	688	2	S55349
15	425.5	7.0	735	2	I50630
16	423	7.0	732	2	S35691
17	412	6.8	787	2	S68699
18	409.5	6.7	645	2	I50680
19	409.5	6.7	663	2	S11521
20	409	6.7	686	1	A44842
21	407	6.7	682	1	JH0560
22	406	6.7	664	2	S11517
23	406	6.7	690	2	A42161
24	401	6.6	691	2	JC6509
25	398.5	6.5	688	2	B42161
26	397	6.5	688	2	S23606
27	397	6.5	857	2	S62694
28	393	6.5	883	2	T07651
29	391.5	6.4	677	2	S32816

30	391.5	6.4	690	1	S07103	CGMP-gated ion cha
31	385	6.3	887	2	T03939	potassium channel
32	384.5	6.3	697	2	H85205	potassium channel
33	384.5	6.3	697	2	T04931	potassium channel
34	382	6.3	828	2	T52046	potassium channel
35	379	6.2	575	2	I59327	olfactory cyclic n
36	375.5	6.2	772	2	S28292	hypothetical prote
37	375.5	6.2	880	2	F85381	potassium channel-
38	375.5	6.2	916	2	T05360	probable potassium
39	367.5	6.0	845	2	T07052	potassium channel
40	352.5	5.8	807	2	T12177	hypothetical prote
41	346	5.7	800	2	T19627	potassium channel
42	342.5	5.6	662	2	T04461	hypothetical prote
43	340.5	5.6	673	2	T20936	hypothetical prote
44	330	5.4	705	2	T06682	hypothetical prote
45	330	5.4	888	2	D84650	probable potassium

ALIGNMENTS

RESULT 1
I38465
probable potassium channel subunit - human
C:Species: Homo sapiens (man)
C:Date: 31-May-1996 #sequence_revision 31-May-1996 #extl_change 05-Nov-1999
C:Accession: I38465
R:Winkler, J.W.; Ganetzky, B.
Proc. Natl. Acad. Sci. U.S.A. 91, 3438-3442, 1994
A:Title: A family of potassium channel genes related to eag in Drosophila and mammals
A:Reference number: A54953; MUID:94211879; PMID:8159766
A:Accession: I38465
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-1159 <RES>
A:Cross-references: EMBL:004270; NID:9487737; PIDN:AAA62473.1; PID:9487738
C:Superfamily: CAMP receptor protein cyclic nucleotide-binding domain homology
F:742-858/Domain: CAMP receptor protein cyclic nucleotide-binding domain homology <CA

Query Match 100.0%; Score 6079; DB 2; Length 1159;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1159; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	MPVRGHAAPONTFLDTIIRKFEKGRKFTIANARVENCAYICNDGFCGLGYSRAEYM	60
DB	1	MPVRGHAAPONTFLDTIIRKFEKGRKFTIANARVENCAYICNDGFCGLGYSRAEYM	60
QY	61	QRPCTCDFLHGPRTORRAAQAIAQLGAEEKVEIAFYRKDGSCTCLVDVVPVKNEDG	120
DB	61	QRPCTCDFLHGPRTORRAAQAIAQLGAEEKVEIAFYRKDGSCTCLVDVVPVKNEDG	120
QY	121	AVIMEFLNFVMEKDMVSGPAHDTHNGPPTSWLAPGRAKTFRKLPAALLALTARESSV	180
DB	121	AVIMEFLNFVMEKDMVSGPAHDTHNGPPTSWLAPGRAKTFRKLPAALLALTARESSV	180
QY	121	AVIMEFLNFVMEKDMVSGPAHDTHNGPPTSWLAPGRAKTFRKLPAALLALTARESSV	180
DB	121	AVIMEFLNFVMEKDMVSGPAHDTHNGPPTSWLAPGRAKTFRKLPAALLALTARESSV	180
QY	181	RSGGAGGACAPAAVVDVLTTPAAPSSESLADEVTAMNHNHAGGPAERARIVGPGSP	240
DB	181	RSGGAGGACAPAAVVDVLTTPAAPSSESLADEVTAMNHNHAGGPAERARIVGPGSP	240
QY	241	PSRAGQOLPSRAHSLNPDASGSSSLATRSRESCASVRRASSADDIEAMRAGVLPAPP	300
DB	241	PSRAGQOLPSRAHSLNPDASGSSSLATRSRESCASVRRASSADDIEAMRAGVLPAPP	300
QY	301	RHAATGAMHPLRSLGSLNSTDVLYRTISKIPQITTLNFVDLKGDFLAPTSDEIITA	360
DB	301	RHAATGAMHPLRSLGSLNSTDVLYRTISKIPQITTLNFVDLKGDFLAPTSDEIITA	360
QY	361	PKIKERTHNVTEKVVQVLSGADVLPDYKQAPRIHRWTLTLYSPFKAWMDLITLLVY	420
DB	361	PKIKERTHNVTEKVVQVLSGADVLPDYKQAPRIHRWTLTLYSPFKAWMDLITLLVY	420
QY	421	TAVFTPYSAFLKETEEGPPATECGYACQAPLAVVDLIVDINFIVDILINFTTYVNANE	480
DB	421	TAVFTPYSAFLKETEEGPPATECGYACQAPLAVVDLIVDINFIVDILINFTTYVNANE	480

Db 421 TAVTPYSAALFKETEEGPATCGYACQPLAVVDLIVIMFIVDILINERTYVANE 480
 QY 481 EVSHPGRIAVHYKFGWFLIDMVAIPEDLLIFSGSSEELIGLKTARLLVRYARKLD 540
 Db 481 EVSHPGRIAVHYKFGWFLIDMVAIPEDLLIFSGSSEELIGLKTARLLVRYARKLD 540
 QY 541 RYSEGAANFLMCTFALLHMLACIWAIGNMEOPHMDSRIGLHNLGDOIGKPYNS 600
 Db 541 RYSEGAANFLMCTFALLHMLACIWAIGNMEOPHMDSRIGLHNLGDOIGKPYNS 600
 QY 601 GLGSPKIDKYVTALYFTFSSITSVGFNVSPTNSEKIFISICVMLIGSLMVAISFGNVS 660
 Db 601 GLGSPKIDKYVTALYFTFSSITSVGFNVSPTNSEKIFISICVMLIGSLMVAISFGNVS 660
 QY 661 AIIORLYSGTARYHTOMLRVREIFRHOIPNPLRQRLREYFQHAMSYNGIDMNAVLKGF 720
 Db 661 AIIORLYSGTARYHTOMLRVREIFRHOIPNPLRQRLREYFQHAMSYNGIDMNAVLKGF 720
 QY 721 PECLQADICLHNLNSLOHCKPFRGATKGLRALAMKFTTHAPPGDTLVHAGDILLTALY 780
 Db 721 PECLQADICLHNLNSLOHCKPFRGATKGLRALAMKFTTHAPPGDTLVHAGDILLTALY 780
 QY 781 FISRGSIEILRGDVVALIGKNDIFGEPLNLYARPKSNGDVRAITYCDLKHHRDDLE 840
 Db 781 FISRGSIEILRGDVVALIGKNDIFGEPLNLYARPKSNGDVRAITYCDLKHHRDDLE 840
 QY 841 VLDNMPERSDHFWSSLEITFNLKNDTMIKPSGSGTELEGFSRQKRKLSFRRTDKOTE 900
 Db 841 VLDNMPERSDHFWSSLEITFNLKNDTMIKPSGSGTELEGFSRQKRKLSFRRTDKOTE 900
 QY 901 QPGEVSALGPRAGAPSSRRRPGGPGWGESPPSSGSESSSEDEDGPGSSSPLRLVPSS 960
 Db 901 QPGEVSALGPRAGAPSSRRRPGGPGWGESPPSSGSESSSEDEDGPGSSSPLRLVPSS 960
 QY 961 PRPGEPPGSGPLMEDCKESSDTCNPLSGAFSGVSNIFSFWDGSRGROYELPRCPAPTP 1020
 Db 961 PRPGEPPGSGPLMEDCKESSDTCNPLSGAFSGVSNIFSFWDGSRGROYELPRCPAPTP 1020
 QY 1021 SLNLNIPSSPGRPRGDEVSRLDALQRLNLETRLSDMATVQLQRLQMTLVPPAYSA 1080
 Db 1021 SLNLNIPSSPGRPRGDEVSRLDALQRLNLETRLSDMATVQLQRLQMTLVPPAYSA 1080
 QY 1081 VTTGCGGTSTSPILPVSPLTLTLDLSQVSOVMACEELPPGAPELPQEGPTRLSLPG 1140
 Db 1081 VTTGCGGTSTSPILPVSPLTLTLDLSQVSOVMACEELPPGAPELPQEGPTRLSLPG 1140
 QY 1141 QLGLATSQPLHRHSGSDPCS 1159
 Db 1141 QLGLATSQPLHRHSGSDPCS 1159

RESULT 2

T19579

hypothetical protein C30D11.1 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999

C:Accession: T19579

R:Morimoto, B. submitted to the EMBL Data Library, August 1994

A:Reference number: Z19145

A:Accession: T19579

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-514 <WIL>

A:Cross-references: EMBL:Z35596; PIDN:CAA84644.1; GSPDB:GN00021; CESP:C30D11.1

A:Experimental source: clone C30D11

C:Genetics:

A:Gene: CESP:C30D11.1

A:Map position: 3

A:Introns: 49/3; 77/2; 151/3; 192/3; 244/2; 264/1; 366/3; 435/2; 479/1

Query Match 29.1%; Score 1767; DB 2; Length 514;

Best Local Similarity 72.6%; Pred. No. 6.6e-104;

Matches 342; Conservative 51; Mismatches 62; Indels 16; Gaps 6;
 QY 376 QVLSGADVLEPYLQAPRIHRWTLHYSPKAVMDMILLVYTAFTPYSAFLKE 435
 Db 49 KVLISGADVLEPYLQAPRIHRWTLHYSPKAVMDMILLVYTAFTPYSAFLKE 108
 QY 436 TEEGPATCGYACQPLAVVDLIVIMFIVDILINERTYVANE--VSHPGRIAVH 492
 Db 109 LQD--TAKSRFTEPLEIVDLIVIMFIVDILINERTYVENEDEACQVSDPGKIATH 165
 QY 493 YFKWMLIDMVAIPEDLLIFSGSSE--LIGLKTARLLRVRVARKIDRYSEGAAY 549
 Db 166 YFKWMLIDMVAIPEDLLIFSGSSE--LIGLKTARLLRVRVARKIDRYSEGAAY 225
 QY 550 LFLMCTFALLHMLACIWAIGNMEOPHMDSRIGLHNLGDOIGKPYNSG--GAGP 605
 Db 226 LFLMCTFALLHMLACIWAIGNMEOPHMDSRIGLHNLGDOIGKPYNSG--GAGP 283
 QY 606 SIKDYVVALYFTFSSITSVGFNVSPTNSEKIFISICVMLIGSLMVAISFGNVAIIOR 665
 Db 284 TLKRYVVALYFTFSSITSVGFNVSPTNSEKIFISICVMLIGSLMVAISFGNVAIIOR 343
 QY 666 LYSGTARYHTOMLRVREIFRHOIPNPLRQRLREYFQHAMSYNGIDMNAVLKGFPCLO 725
 Db 344 LYSGTARYHTOMLRVREIFRHOIPNPLRQRLREYFQHAMSYNGIDMNAVLKGFPCLO 403
 QY 726 ADICLHNLNSLOHCKPFRGATKGLRALAMKFTTHAPPGDTLVHAGDILLTALYFTSRG 785
 Db 404 ADICLHNLNSLOHCKPFRGATKGLRALAMKFTTHAPPGDTLVHAGDILLTALYFTSRG 463
 QY 786 STEILRGD-VVVALIGKNDIFGEPLNLYARPKSNGDVRAITYCDLKHHR 835
 Db 464 STEILRGD-VVVALIGKNDIFGEPLNLYARPKSNGDVRAITYCDLKHHR 514

RESULT 3

T31354

probable potassium channel elk chain 1 - rat

C:Species: Rattus norvegicus (Norway rat)

C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 29-Oct-1999

C:Accession: T31354

R:Engelard, B.; Neu, A.; Ludwig, J.; Roepert, J.; Pongs, O.

A:Submitted to the EMBL Data Library, July 1998

A:Description: Identification of three rat potassium channel genes homologous to D. m

A:Reference number: Z20983

A:Accession: T31354

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-1017 <ENG>

A:Cross-references: EMBL:A1007628; NID:e1329997; PID:e1329998; PIDN:CAA07587.1

A:Experimental source: cortex

C:Genetics:

A:Gene: elk1

C:Keywords: potassium channel

Query Match

Best Local Similarity

Matches 417; Conservative 172; Mismatches 341; Indels 312; Gaps 36;

QY 1 MPVARGHVAONFPLDITIRKFEQSKRTIINAR-VENCAVICYNGRELCGYSAEV 59
 Db 1 MPVARGHVAONFPLDITIRKFEQSKRTIINAR-VENCAVICYNGRELCGYSAEV 60
 QY 60 MOPRCTDPLHGPTORRAAQIAQLGAEERKVEIAYFRKDGSCFLVADVPAVNE 119
 Db 61 MOKTCSGRFLXGPTSPALQRLQKALEGHOERAEICFRKXGSAWCLDMPIKREL 120
 QY 120 GAVIMFTLNFVYMEKDMVSGPAHDITNHRGPPYSWLAAGRAKFRKLALALTARESS 179
 Db 121 GEVVLFLFSK-----D 132
 QY 180 VRSGAGAGAPGAVVVDVLTTPAASSESLADEVTAMDNHVAIGCPAERRALVPGS 239
 Db 180 VRSGAGAGAPGAVVVDVLTTPAASSESLADEVTAMDNHVAIGCPAERRALVPGS 239

```
Db 133 ISGSGPGLGSPG---IHGD-----NNNHNSLIG---RR-----160
QY 240 PRSAPGQLPSPRAHSLNPDAGSSCSGLARTRESCASVRRASADIEAMAGVLPPE 299
Db 161 -----GASSRLNSTRON-----173
QY 300 PRHASTGAMHPLKSGLLNSTSDSLVRYRTISKIPQITLNFVDLKGPPLASPTSDREII 359
Db 174 -----RTV--LHRLTGHF-----GRDQGSVKANSNWF 199
QY 360 APRKIKRTNHTVTEKTVQVLSLADVLPYKQLAPRIHRWTLIHSYSPKAVMDLILLVI 419
Db 200 EPK-----PSVPEYKVASVSGSCLLILHSIPKAVMDLILLVIATF 239
QY 420 YTAVFPTPSAFLFKETEGEPATECGYACQPLAVVDLYDIMEFYDILINFRTYVNNAN 479
Db 240 YVAVIYPIYVCF---AGDDTPIYS-----RHTLVSDIAVEMLEFIDILINFRTYVSOS 291
QY 480 EEVYSHPRGRIAVHYFKGWFLLDMVAIIPEDLT-IFGSGSEELIGLKTARLLRLVAVARK 538
Db 292 GGVVSAAPRSIGLHYLATWFFVDLIAALPDLVYFNITVSLVHLKTVRLRLRLK 351
QY 539 LDRYSEGAAYFLMCTPALLAHMLACTIWAIGNEDPHMS---RIGMLNLDQIGK 595
Db 352 LERYSQCASAVVLTLLMSVALLAHMMAWYVIGREMEANDPLMDIGLHLEGRLEE 411
QY 596 PYSSGLGSPSIXDKVYVTLALYFESSLTSGVGNVSPNTSEKIFSCVWLIGSLMVASI 655
Db 412 PYVNGSAGPBRSAVIAALYFTLSSLTSGVGNVSPNTDAEKIFSCVWLIGALMVAAY 471
QY 656 FGNVSAIIORLYSGTARHYTOMLRVREFIRHOIIPRLRQLEEYFOHAMSNTGIDMNA 715
Db 472 FGVNTALIORMYSRSLYHSRMDKDFIRVHRLPRLQOMLEYQOTTWAVNSGIDANE 531
QY 716 VLAKEPCELOADICHLNLSLOHCKPFRGATKGCRLAAMKKTTHAPPGDTLVHAGDL 775
Db 532 LKRPDEDELRAIDAMHINREILO-LPLFGASRCGLRALSIHKTSECAEPFELLRGDA 590
QY 776 LPLALYISRSIEILGADVVALTGKNDIFGEPL-NLYARPG-----KSNDDVRLATY 827
Db 591 LQAHYVYSGSLELNDVTLALTGKDLGADIPELGDEPGAGACVLTASDVRLATY 650
QY 828 CDLAKIHRDLEVLDMYEFSDHFWSSL-EITFNLRTNMIIPGSPGSELEGGFSR-Q 884
Db 651 CGLOQLSSRGLAEVLLPYEYVAAFRAGLPRODLTFMLRO-----GSENN-----GLCRFS 700
QY 885 RKRKLSFRRR---TDKDTQEGEVALGPRAGAPSSRG-----RPGGFWG-----928
Db 701 RSPRLSQARSDDLGSSSDKTLPISITTEGMEPGAGSKPRRPLLNLSPARRGSLVSL 760
QY 929 ---ESP--SSGSPSPESSEDECP---GRSSP-LRLVPPSP---RRP-----GEP 967
Db 761 LGEELPFPALVSPSPSPALAGRSSPPLHGPBGSAAMKBPQOLLTPPLGTFGP 820
QY 968 PGGEPIMEDCEKSDTCDNPLSGAFSGVSNIFSWGDSRGROYELPRCAPPTSLNINPL 1027
Db 821 DLSFMTVVDGEDSSNTAAEATFRFS-----KREPITTRSOAPL 859
QY 1028 SSP-----GRPRGVESTRDALQLOLNLE---TRLSADMATVLOLLQORMTLVPPAY 1078
Db 860 SGRLRELATEAEAEVEKVKVCRILNOEISRLNOEVSOLRELROVGLLQAR--LGPSSH 917
QY 1079 SAATY-----PGGPTSTPLPVSLPPLTILDSLSQVSOQMAC 1117
Db 918 PPDSTMLPDLPCPHQRPCTISPMHSGPPPLQMTTLLAVVHCPASVGTVEIGATPSE-LRS 976
QY 1118 EELPQAPALPQEGPTRRLSLPGOLALTSQPLHRHGSDPGS 1159
Db 977 SMVPPPPSEPPDLGFS-----PYPEASPLTSLIKHSFQSGS 1013
```

RESULT 4
T17367

```
potassium channel protein elki - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C:Accession: T17367
R:Shi, W.; Wang, H.S.; Pan, Z.; Wymore, R.; Cohen, I.S.; McKinnon, D.; Dixon, J.E.
J. Physiol. 511, 675-682, 1998
A:Title: Cloning of a mammalian elk potassium channel gene and EAG mRNA distribution
A:Reference number: 218731; MVID:98382545; PMID:9714851
A:Accession: T17367
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-1102 <SH1>
A:Cross-references: EMBL:AF061957; NID:93659689; PID:93659690; PIDN:AAC61520.1
C:Genetics:
A:Gene: elki
C:Function:
A:Description: may play a role in the sympathetic nervous system
C:Keywords: potassium channel

Query Match 25.5%; Score 1551; DB 2; Length 1102;
Best Local Similarity 32.6%; Pred. No. 7.9e-90;
Matches 383; Conservative 169; Mismatches 283; Indels 340; Gaps 30;

QY 1 MPVRGHVAPQNTFLDIIIRKFGQSRKFLIANARY-ENCAYITCNDGFCGLGYSRAEV 59
Db 1 MPVKGGLAPQNTFLDIIATRFDTGHSNFIANAQYAKGPIYVCSGDFCELAGFARTEV 60
QY 60 MQRPCIDPLHGRTRORRAAOIAQALLGAEEKVELATFRKQSGFCLIVDVPYKNEI 119
Db 61 MOKSCCKFLFGEVTEOIMLOIETKSLSEKVEKFEKEMFKKGAQGFVCLLDIVPIKNEK 120
QY 120 GAVIMTFLNEVMEKDMVSGPAHDNHRPPTSWMLAPGAKTFRLKPLALLTARESS 179
Db 121 GDVYLFASF---KDIIDTKV-----138
QY 180 VRSGAGGAGAPAVVVDVLTLPAPSSSLADEVTAMDNHVAGLPAPERALVPGS 239
Db 139 -----KITSDEK-----KEDR-----149
QY 240 PRSAPGQLPSPRAHSLNPDAGSSCSGLARTRESCASVRRASADIEAMAGVLPPE 299
Db 150 ---AKRG-----SRAGSHPSARRSR-----168
QY 300 PRHASTGAMHPLKSGLLNSTSDSLVRYRTISKIPQITLNFVDLKGPPLASPTSDREII 359
Db 169 -----AVLYHISGLDREKKKLINNVFVDPK-----197
QY 360 APRKIKRTNHTVTEKTVQVLSLADVLPYKQLAPRIHRWTLIHSYSPKAVMDLILLVI 419
Db 198 -----AFPEYVSDAKSKSLIHLHSTFKAGMDWLLIATF 233
QY 420 YTAVFPTPSAFLFKETEGEPATECGYACQPLAVVDLYDIMEFYDILINFRTYVNNAN 479
Db 234 YVAVIYPIYVCF---AGDDTPIYS-----RHTLVSDIAVEMLEFIDILINFRTYVSOS 291
QY 480 EEVYSHPRGRIAVHYFKGWFLLDMVAIIPEDLT-IFGSGSEELIGLKTARLLRLVAVARK 538
Db 285 GOVIFEARSTCIHYVTTFEIDILALPDLTAFVNTVAVSLVHLKTVRLRLRLK 344
QY 539 LDRYSEGAAYFLMCTPALLAHMLACTIWAIGNEDPHMS---RIGMLNLDQIGK 595
Db 345 LDRYSQHSHTVLLNLSMFLAHMMACTIWAIGKKEKRENSILKMEVGLHLEGRLES 404
QY 596 P-YNSGGLGSPSIXDKVYVTLALYFESSLTSGVGNVSPNTSEKIFSCVWLIGSLMVAS 654
Db 405 PYVNGTTLGSPSIRSAVIAALYFTLSSLTSGVGNVSPNTDAEKIFSCVWLIGALMVAAY 471
QY 656 FGNVSAIIORLYSGTARHYTOMLRVREFIRHOIIPRLRQLEEYFOHAMSNTGIDMNA 715
Db 465 VFGNVTAIIORMYSRSLYHSRMDKDFIRVHRLPRLQOMLEYQOTTWAVNSGIDANE 531
QY 716 VLAKEPCELOADICHLNLSLOHCKPFRGATKGCRLAAMKKTTHAPPGDTLVHAGDL 775
Db 715 AVLAKEPCELOADICHLNLSLOHCKPFRGATKGCRLAAMKKTTHAPPGDTLVHAGDL 774
```

Dh 525 ELKDFPEDELRSDDTMMHINKNEIIQ-LSLFECASRGCLRSLSLAKTSTFCACGEYLLROGD 583
Qy 775 LFLALYISKGSIEILRGDVVAIIIGKNDIEGEPLENTYARPGKNGDVRALTYCDLKHII 834
Dh 584 ALQAIYVCGSGMELVLDMSVLAIIIGKNDIEGEPLENTYARPGKNGDVRALTYCDLKHII 643
Qy 835 RDLLEVLDMYPERSDHFWSSL-EITFNLD-----TNMI PGSGSTEEGG 880
Dh 644 LKGLFEVLGYPEYAHKFVEDIQHDLTYNLREGHSDVLSRLSKSTVP-----QAEPRGN 699
Qy 881 FSRKRKKLSRRRTDQTDQPGVSAALGPRACAGSSRGPRGPGMGESPPSGSPSPS 940
Dh 700 GS-1KKRLPSLVEDEEEVEEETSLSP-----ITYRG-----SSVSH 738
Qy 941 SEDEGPRSSPLRL-----VPFSSP-----RPPGEP----- 968
Dh 739 SKRTGSSKSYIGLSLKQLTSGTVFHSPIRVSSANSRKTQEOADPPHNGTKENKLVOL 798
Qy 969 -----GGEPLM-----EDCEKSDTCNPLSGAFSGVSNIFSPWDSRGRQYQELPRC 1015
Dh 799 CSLGTAGPSPSPRIYDGIEDGNSSEET-----QTFPGSE-----QIRPEPRI 842
Qy 1016 PAPPSLNLPLSSPGRRPGDVE-----SRDLAQRLKLE-----TRISA 1058
Dh 843 ---SPSL-----GESEIGAFLEFIKAETTKQOINKLNSEVTLTQESOLGK 886
Qy 1059 DMATVLOLOROMTLVPPAYSAVTTPGPTSTSP 1093
Dh 887 DMRSIMOLENITLSPQPSORC-----SLHPTSTCP 917

RESULT 5
T31100
probable potassium channel 2 - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 22-Oct-1999
C:Accession: T31100
R:Engelard, B.; Neu, A.; Ludwig, J.; Roepert, J.; Pongs, O.
Submitted to the EMBL Data Library, July 1998
A:Description: Identification of three rat potassium channel genes homologous to D. melanogaster
A:Reference number: Z20983
A:Accession: T31100
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-1087 <EMBL>
A:Cross-references: EMBL:AJ007627; NID:e1329995; PID:e1329996; PIDN:CAA07586.1
A:Experimental source: cortex
C:Genetics:
A:Gene: elk2
C:Keywords: potassium channel

Query Match 25.3%, Score 1537.5; DB 2; Length 1087;
Best local similarity 32.4%, Pred. No. 5.5e-89;
Matches 419; Conservative 143; Mismatches 357; Indels 373; Gaps 35;

Qy 1 MPVRGHVAPONTFLDTIRKFEQSRKFIANARVENC-AVIYCNDFCELCYSRAEV 59
Dh 1 MPVRGHVAPONTFLDTIRKFEQSRKFIANARVENC-AVIYCNDFCELCYSRAEV 60
Qy 60 MQRCTCDLHPTORTARRAAQIAOALLGAERKVEVLAFTKDDSCFLCLVDVVPKNE 119
Dh 61 MQRCTCDLHPTORTARRAAQIAOALLGAERKVEVLAFTKDDSCFLCLVDVVPKNE 120
Qy 120 GAVIMFLINFEVMEKMWSPANDTNHGRPTSMIAPGRAKFTRLKIPALLALTARSS 179
Dh 121 GAVIMFLINFEVMEKMWSPANDTNHGRPTSMIAPGRAKFTRLKIPALLALTARSS 153
Qy 180 VRSAGAGAGARGAUVVVDVLTTPAPSSSESLALDEVTAMDNHVAGLGAEEERRALVPGS 239
Dh 154 -RYGRAGSKG----- 162
Qy 240 PPRSAPOQLSPRAHSLNPDASGSSCLARTRSRESCASVRRASADDIEMRAGVLP 299

Dh 163 -----FNANRRRSNAVLYHL-----SGHLQK 184
Qy 300 PRHASTGAMPLRLRSGLNSTSDSDLVRYRTISKIPQITLNFVDLKGDPFLASPTSDREII 359
Dh 185 PK-----GKHKLNGVGEKPN----- 201
Qy 360 APKIKERTHNATEKVTVLSGLADVLPDYKIQAPRIHRWTILAHSPKAWMDLILLVYI 419
Dh 202 -----LPEYKAAIRKSPFILLHCGALRAVWDGFIILATL 236
Qy 420 YTAFTPYSAFLKTEDEGPRATECGYACQAPLAVVDLIYDIMVILVILNFTTYNAN 479
Dh 237 YTAFTPYSAFLKTEDEGPRATECGYACQAPLAVVDLIYDIMVILVILNFTTYNAN 288
Qy 480 EGVVSHGRVIAVHYFKGMFLIDVAAIPFDLL-IFGSGSEELGLLTARLRLVYARK 538
Dh 289 GQVYFAKSLICHLVTTWFLDVAIALPFDLLHFKVNVYGAHLKLTVRLRLRLPLR 348
Qy 539 LDRYSEGAVALFLMCTFALIAHMLACIWAYIGNMEQPHMD--RIGWLNIGDQIGK 595
Dh 349 LDRYSEGAVALFLMCTFALIAHMLACIWAYIGNMEQPHMD--RIGWLNIGDQIGK 408
Qy 596 PY-----NSSG-----LGPSIKDKVYVALYFTFSSLT 624
Dh 409 PYLVYLSRSPDGGNSSGSENCSSGSGSEANGTLELLGSPSLSAVITSLYFALSLS 468
Qy 625 VGFNVSPNTNSEKIFISICVWLISLWYASIFGNVSAIIOFLYSGTARYHTOMLRVEFI 684
Dh 469 VGFNVSPNTNSEKIFISICVWLISLWYASIFGNVSAIIOFLYSGTARYHTOMLRVEFI 528
Qy 685 RHOIIPNLRORLEEVYOHAWSYNGIDMAVYLGPEPCLOADICLHNSLLOHCKPFR 744
Dh 529 RHOIIPNLRORLEEVYOHAWSYNGIDMAVYLGPEPCLOADICLHNSLLOHCKPFR 587
Qy 745 GATKGCCLRALAMKKTTHARPDDTLVHAGDILTLYFISNGSTILGDDVVAILIGKNDI 804
Dh 588 AASRGCCLRALAMKKTTHARPDDTLVHAGDILTLYFISNGSTILGDDVVAILIGKNDI 647
Qy 805 FGEPLNLYARPGKNGDVRALTYCDLKHIIHDDLEVLDMYPERSDHFWSSL-EITFNLD 862
Dh 648 FGEPLNLYARPGKNGDVRALTYCDLKHIIHDDLEVLDMYPERSDHFWSSL-EITFNLD 707
Qy 863 R-----DTNMI PG-----SPGSTEEGG 880
Dh 708 GAGVSAEVDTSLSGDNLTLMSTLEKETDGEQHTISPADEPSSPLSPGCT----- 762
Qy 881 FSRKRKKLSRRRTDQTDQPGVSAALGPRACAGSSRGPRGPGMGESPPSGSPSPS 940
Dh 763 -SSSSAAKLSPKRT---APRPLRGGRPRRAGVLRKPGSPSAHPRTLDSGLRLPMPWN 818
Qy 929 ESPSGSPSPSPSED-----EGPGRSSPLRLVFPSSPRPPEPPG--GE 971
Dh 819 ESPSGSPSPSPSED-----EGPGRSSPLRLVFPSSPRPPEPPG--GE 868
Qy 972 PLMEDCEKSDTCNPLSGAFSGVSNIFSPWDSRGRQYQELPRCAPPSLNLPLSSPG 1021
Dh 869 PLMEDCEKSDTCNPLSGAFSGVSNIFSPWDSRGRQYQELPRCAPPSLNLPLSSPG 919
Qy 1032 RRRP-----GDVESRLDLQRLNRLFTRLSDAMATVLIQILQRLMTLVPPAYSAY--TPRG 1085
Dh 920 RRRP-----GDVESRLDLQRLNRLFTRLSDAMATVLIQILQRLMTLVPPAYSAY--TPRG 966
Qy 1086 PGPTSTPDL-----PVSPDL-----TLTDLSLQVSOFMACEELPRGAPELPQ 1129
Dh 967 PGPTSTPDL-----PVSPDL-----TLTDLSLQVSOFMACEELPRGAPELPQ 1026
Qy 1130 EGPTTRLSLPGQLGALTS--QPLRHGSDPGS 1159
Dh 1027 EGARTGTAPARYSQAEATSTGP-----PPGS 1052

RESULT 6
T13168

probable potassium channel elk chain - fruit fly (Drosophila melanogaster)
C:Species: Drosophila melanogaster
C:Date: 13-Aug-1999 #sequence_revision 13-Aug-1999 #text_change 17-Nov-2000
C:Accession: T13168
R:Wanke, J.W.; Ganetzky, B.
Proc. Natl. Acad. Sci. U.S.A. 91, 3438-3442, 1994
A:Title: A family of potassium channel genes related to eag in Drosophila and mammals.
A:Reference number: A54953; MUID:94211879; PMID:8159766
A:Accession: T13168
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-1284 <MAX>
A:Cross-references: EMBL:U04246; NID:g487735; PID:g487736; PIDN:AAA62472.1
C:Genetics:
A:Gene: elk
A:Cross-references: FlyBase:FBgn0011589
A:Map position: 2R

Query Match 24.4%; Score 1485.5; DB 2; Length 1284;
Best Local Similarity 32.3%; Pred. No. 1.3e-85;
Matches 366; Conservative 173; Mismatches 330; Indels 265; Gaps 27;

QY 1 MEYRKHVAPONTFLDTITRKFEQGRKTIANARVENCAYIYCNDFCELCYSRAEYV 60
DB 1 MPARKGLAPONTFLDTITRKFEQGRKTIANARVENCAYIYCNDFCELCYSRAEYV 60
QY 61 ORPCTDPLHGPRTORRAAOIAQALGAERKVEIAFYRKOSCFCLVDVVPVKNEDG 120
DB 61 QKQCSHFLEPPTKEKHQOIEKLSNMELKLEYIFKKEBAPWCPLFDIYPIKNER 120
QY 121 AVIMFLNFEVMEKDVSSPAHDTNHRPPTSWLAPGRAPKTRFLPALALTAARESSV 180
DB 121 DVVLFASHKIDITHTKMLEMNVNE-----ECDSVFALTAALLGA 159
QY 181 RSGGAGCAAPGAVVVDVLTTPAAPSSESLADEVTAMDNHVAAGLCPAERRALVPGSP 240
DB 160 RFRAGSNAG-----MLGIG-----GLPGLGSP 181
QY 241 PRSAPQLPSRAHSLNPDASGSSCLARTSRSCASVRRASASADIEAMRAGVLP 300
DB 182 AAS-DQDTEAGEGNNLDVPA---GCMGRRSR-----AVL----- 213
QY 301 RHAASGAMHPLNSGLINSTSDSLVRYRTISKIPQITLNFVDLKGDPFLASPTSDREIIA 360
DB 214 -YQLSGHYKPEKGV-----KTKKLGN---NFMHSTEAPF----- 245
QY 361 PKIKETHNVEKYVYOVLSGADVLPYKLOAPRIHRTITLHSPKAVMDLILLVY 420
DB 246 -----PEYKTOSIKSRLLILPHYGEKGIWDVILVATPEY 280
QY 421 TAVFTPYSAFLKTEEGRPATEGACOPLAVVLDIYDIMEITVILNFTTYVNAME 480
DB 281 VALMYYVNAF-----AKADQRTKVSQVYIEALTYVILNFKTTFEYSKRG 326
QY 481 EYVSHBGRIAVHYFKGMPFLIDVVAALIPDLLJFG---SGSEELIGLLKTLARLVRAR 537
DB 327 EYVSNKQALINYLRCGFALDILALPFDHLVASDLYDGDSDIHLVKTLRLRLARLDQ 386
QY 538 KIDRYEYGAAYFLMLCTFALIAHLACIWAIGMEOPHM-DSKIGMLHNIQDQIGRP 596
DB 387 KIDRYSOHTAMITLTLMFSEFTLAHMLACIWAYIAKEYEWPEPESINIGQLAER---- 442
QY 597 YNNSGIGSPSIKDKYVYATYFSSLSYGFVGNVSPNTSEKJFSCVMILGSLMAYSTF 656
DB 443 -KNASVAILTAEYVSTALYFTTSLTSYGFVGNVSNNTAEKYFTIIMLIGLMAVAVF 501
QY 657 GNVSAITQRLYSGTARYHTQMLRVREFIRPHQIPNDLQRLLEEFQOHANSTYNGIDMANV 716
DB 502 GNTVATIQMYSRSLYSESKWMDLKDFVALHNPKEKRIEDYFQTSMSLSHGIDYET 561
QY 717 LKGFPCLOADICLHNBSLQCHKRFKATKGCCLALAMKFTTTPARPQDTLVHAGDLL 776
DB 562 LRPEPELKGVDVSMHLHREITLQ-LPIFEAASQCLRLSLHKTNCFACGEYVLIHKGDAL 620

QY 777 TALYFIRSGSIELRGDVVAALIGKNDIFGEPLNLYARPKSNG----- 820
DB 621 NTVIYLCNMSMEVITKNDMVAALIGKNDIVGSDINVL-VATNSGQMTATNTNSAGDVYV 679
QY 821 ---DVRLATYCDLHKTRDLEVLDMYPEFSNHEWSSL--ETTFNLR-----DTNMT 869
DB 680 SSSDIALYICDCKICIMGVLVEVLRLYPEYQOOFANDIONHDTLCMLREYENODSDIG 739
QY 870 GSPGSELELGGFGRQKRKLSFRRRTDKDTEQGEVSALGPAGAGSPSSRGPGPWE 929
DB 740 SEPLPSTISE-----DDEKREAEKGGKEKENGSGPSPGASPLNINIS 782
QY 930 SPSSGSPSESSSEDEGRSSSPRLYFPSSPPRPPEPGGEPL--MEDCEKSDTCNPL 987
DB 783 SPLHATRSF-----LLGMSPRNQALHQRGRSLITLRETKNKRHLNMA 826
QY 988 SGAFSGVSNIFSWGDSRGROYELPRCAPRPSLNLPLSPGRRRPQDVESRLDALOR 1047
DB 827 CSLDRG-----SFEK-----PEPLEEQQSGCKRPSLE---RLDSQVS 861
QY 1048 QLNRLTRLSADMATYVQLQRLQMTLV-----PRASVATTTGPGPT 1089
DB 862 TLHQDVAAQDSAEVRNALISALQ-EMTETSNAMTSHSLKPPARPSPINISGACT 914

RESULT 7
A40853
potassium channel protein eag - fruit fly (Drosophila melanogaster)
C:Species: Drosophila melanogaster
C:Date: 10-Jul-1992 #sequence_revision 10-Jul-1992 #text_change 24-Sep-1998
C:Accession: A40853
R:Wanke, J.; Drysdale, R.; Ganetzky, B.
Science 252, 1560-1562, 1991
A:Title: A distinct potassium channel polypeptide encoded by the Drosophila eag locus
A:Reference number: A40853; MUID:91262635; PMID:1940699
A:Accession: A40853
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-1174 <MAX>
A:Cross-references: GB:M61157; NID:g157311; PID:g157312
C:Genetics:
A:Gene: FlyBase:eag
A:Cross-references: FlyBase:FBgn0000535
C:Keywords: transmembrane protein

Query Match 23.8%; Score 1448.5; DB 2; Length 1174;
Best Local Similarity 30.4%; Pred. No. 2.6e-83;
Matches 374; Conservative 174; Mismatches 328; Indels 355; Gaps 33;

QY 4 RGHVAPONTFLDTITRKFEQGRKTIANARVENCAYIYCNDFCELCYSRAEYV 62
DB 5 RGLVAPONTFLDNITRNSQSDSSFTLANQIYDPPYIYCNEFSCKISGIRAEYVOK 64
QY 63 PC--TCDFLHGPRTORRAAOIAQALGAERKVEIAFYRKDG---SCEFL----- 107
DB 65 SCRYVCGFMKGETLKEVGRLEYLENQDQDFEILYKKNNLQCGCALSQGKAQOTGE 124
QY 108 ----CLVDVVPKNEGAVIYMTLNEVMEKDVSSPAHDTNHRPPTSWLAPGRAPKTF 163
DB 125 TPLMLLOVAPITNEDLVVFL-----TF 150
QY 164 RLKLPAALALTAARESSVRSAGAGAPGAVVVDVLTTPAAPSSESLADEVTAMDNHVA 223
DB 151 R-----DITALKQPT- 160
QY 224 GLGPAERRALVPGSPRSAPQQLPSRAHSLNPDASGSSCLAR--TRSRSCASVRR 281
DB 161 ---DSBDTGVLG-----LSKFAKLARSVTRSRQFSANL--- 191
QY 282 ASSADIEAMRAGVLPFRPHASTGAMHPLRSGLNSTSDSLVRYRTISKIPQITLNV 341
DB 192 ----- 191

```

0Y 342 DLKDDPLASTSREIIRAPKIRERTHNVEKYOVLSDGADVLPBVKLOAPRIHRHTIL 401
Db 192 -----PLTKDPTKO-----SNLHMSLSLADIMPQYKQEPKTPPHILL 230
0Y 402 HYSEPKAVMDLILLLVYTAFTPTYSAAFLKETEETEEPPATEGECYACQPLAVDILJVDI 461
Db 231 HYCAFLAIMDVILLCFLPFTYAMVPYNVAFNKKSJED-----VSLLYVDLSIVD 279
0Y 462 MEIYDILINFETTYVYANEEVYSHPGRAVHYFEGWFLIDKVAALIPDLDLIFGS----- 515
Db 280 IFFDIYVINFETTYVGPGEVYVSDPKVIRMYLYKSWFLIIDLSCILPYD--VFNAFDRDED 337
0Y 516 GSEELIGLKARLILAEVAVARKDLRYSEYGAAYFLMLCTFALIAHACIWAIGNME 575
Db 338 GIGSLFSLAKYVRLRLGRVYRKIDRYLEYGAAMLILLCYMYLVANHACIWSIGRSD 397
0Y 576 QPHMDSRIGWLHNGDQIGKPY-----NSSG---LGGPSIKDKYVYALYFTFSSLSVGF 627
Db 398 ADN-GIOXSMWKLANVYQSPSYIWSNDTPPELVNCGSRKSMVYALYFTMTCTSYGF 456
0Y 628 GNSPNTNSEKIFISICVYLIGSLMYASIFGVNSVALIORLXSGTARYTHOMRAREEFLRFH 687
Db 457 GNVAAEIDNEVEYFTCMKIIAALXYATIFGVYTTIIIDQMISATIKYHDMLNWNEFKLH 516
0Y 688 QIPNPLROLEEFYFOHANSYTINGIDMNAVYKGFPECIQADICLHNLNSLLQHCPRFGAT 747
Db 517 EVPRALSERYADVYVSTWAMTKGDTKEVLCNCKRMDKADICVHLNRKVEDEHPTFLRAS 576
0Y 748 KGCIRALAMRKKTTHAPPGDTLVHAGDLTALTYATISGSTEILLGADVVALATKNDIIFE 807
Db 577 DGCIRALAMHFMHSHAPDGLYHTGESIDSLCTYVGSLEVIDDDEVALLIGADVFGD 636
0Y 808 PLANTYARPGKSGNDVRALTYCDLKHIRDDLEVLVDWYPERSDHFWSELETETFMRLDTNM 867
Db 637 QFMKDSAAVGSAANVRAITYCDLHAIKDKLLEVLDFYSAPANSFARNLVLYTNLHR-RL 695
0Y 868 IPGSRGSELEGGTSRQK-----RLTSRRRTDKRTEQGEVYSALGPRGA- 913
Db 696 IFRRVADYKRKEKLEAERKKNPEQLPQNDHLVRIFSKFRPTPOVQAGSKELVGGSSQSD 755
0Y 914 ---GAGPSSRR--PGGP-----WGESPPSGSPSESSDEDEPPGRSS 951
Db 756 VEKDGGEYERKVLPKAPKLOASQATLARQDTIDEGGEVYDSSPPSRSRVYIEBANYSSA 815
0Y 952 PLRLVPSSPPPP-----GEPGGEPLMEDCKSS---DTCNPLSGAQS 992
Db 816 TV-----GPBPATYTTSSAAAGAGVSGGPGSGGTVAALVYTKAQRNLALEREQIEMASS 869
0Y 993 GVSNIFFSWGSRORQVOELPRCAPPPSILNITPLSSGRRPRQDVSRLDALORLNRL 1052
Db 870 RATT-----SDTYTGTLEKTEP-----PTLAQRDLVATVLDKMKADVRLLELRMOORIGRI 918
0Y 1053 ETRLSDAMVATVLOLOROMLVPPA-----YSAVTTTGP-----GPTSTSP 1093
Db 919 EDLIG-----ELVKR---LAPGASSGGMAPDNSSGQTPPGDIBELCAGCAGCGGCTPTTQ 968
0Y 1094 LLPVSPPLTYLDSLDSOVSPMACEELPPGA 1124
Db 969 APPTSAN-TSPDVTIVTIS-----PGA 990

```

RESULT 8
I53197
potassium channel subunit - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 29-May-1998 #sequence_revision 29-May-1998 #text_change 05-Nov-1999
C:Accession: I53197
R:Ludwig, J.; Terlau, H.; Wunder, F.; Bruggemann, A.; Pardo, L.A.; Marguardt, A.; Stühmer
EMBO J. 13, 4451-4458, 1994
A:Title: Functional expression of a rat homologue of the voltage gated ether a go-go po
ts mammalian counterpart.
Reference number: I53197; MUID:95009946; PMID:7925287

QY	Accession:153197	A:Status: preliminary: translated from GB/EMBL/DBJ	QY	Accession:153197	A:Status: preliminary: translated from GB/EMBL/DBJ
DB	A:Molecule type: mRNA	A:Residues: 1-962 <RES>	DB	A:Molecule type: mRNA	A:Residues: 1-962 <RES>
	A:Cross-references: EMBL:Z34264; NID:g557264; PIDN:CAA84018.1; PID:g557265			A:Cross-references: EMBL:Z34264; NID:g557264; PIDN:CAA84018.1; PID:g557265	
Query Match	23.2%: Score 1409.5; DB 2; Length 962;		Query Match	23.2%: Score 1409.5; DB 2; Length 962;	
Best Local Similarity	30.9%: Pred. No.5.7e-81;		Best Local Similarity	30.9%: Pred. No.5.7e-81;	
Matches	338; Conservative 151; Mismatches 275; Indels 329; Gaps 23		Matches	338; Conservative 151; Mismatches 275; Indels 329; Gaps 23	
QY	4 RGHVAPONTFLDTIRKEGOSRFEIIIANAVENCAYIYCNDFGCELGSRAYEMORP	63	QY	4 RGHVAPONTFLDTIRKEGOSRFEIIIANAVENCAYIYCNDFGCELGSRAYEMORP	63
DB	7 RRGVLAAPONTFLENIVRR--SNDTFNVLGNAGIYDMPITYNSDGCCKLSGHRAYEMQMS	64	DB	7 RRGVLAAPONTFLENIVRR--SNDTFNVLGNAGIYDMPITYNSDGCCKLSGHRAYEMQMS	64
QY	64 CTCDFLHGPRIQRRRAAQIQAOLALCAEERKVEIATYRKDSCFCLLDVVPYKPNEDGAVI	123	QY	64 CTCDFLHGPRIQRRRAAQIQAOLALCAEERKVEIATYRKDSCFCLLDVVPYKPNEDGAVI	123
DB	65 SACSFMYGELTDKDKVEKVRQTFENYEMNSPEILMYKKKRTFVPWFVKIAPIRNODKVV	124	DB	65 SACSFMYGELTDKDKVEKVRQTFENYEMNSPEILMYKKKRTFVPWFVKIAPIRNODKVV	124
QY	124 MFLNFEVYMEKDMVGSAPAHDTNHRGPTSWIAPCRATKFRLLKLLALLATARESSVRSG	183	QY	124 MFLNFEVYMEKDMVGSAPAHDTNHRGPTSWIAPCRATKFRLLKLLALLATARESSVRSG	183
DB	125 LFLCTFS-----	131	DB	125 LFLCTFS-----	131
QY	184 GAGGAGAGCAVVDVDTLPAAFSSESLALDEVTAMDNHVALGP-AEERRALVGGSPPR	242	QY	184 GAGGAGAGCAVVDVDTLPAAFSSESLALDEVTAMDNHVALGP-AEERRALVGGSPPR	242
DB	132 -----DITAFKQPIE-----DDSCKGKGFARLRAL-----	158	DB	132 -----DITAFKQPIE-----DDSCKGKGFARLRAL-----	158
QY	243 SAPGOLPBPRAHSLNPDAAGSSCSILARTRSRESCASVRRASSADDIEAMRACVLPPEPRH	302	QY	243 SAPGOLPBPRAHSLNPDAAGSSCSILARTRSRESCASVRRASSADDIEAMRACVLPPEPRH	302
DB	159 -----TSSR-----GVL-----	165	DB	159 -----TSSR-----GVL-----	165
QY	303 ASTGAMHRLRSLGSLNSTDSDLVRYRTISKIPQITLNLNVDLKGDPFLASPTSDRETIAPK	362	QY	303 ASTGAMHRLRSLGSLNSTDSDLVRYRTISKIPQITLNLNVDLKGDPFLASPTSDRETIAPK	362
DB	166 -----QOLAPS 171		DB	166 -----QOLAPS 171	
QY	363 IK--ERTHNVEKTYQVLSGADVLPEYKLAQPIRHRWTILHYSFKAWMDLILLVY	420	QY	363 IK--ERTHNVEKTYQVLSGADVLPEYKLAQPIRHRWTILHYSFKAWMDLILLVY	420
DB	172 YQKGENVHK-HSRLEAVNLQSGDILPQYKQEARKTPRHILHYCYFKTQWDMITILIRPY	230	DB	172 YQKGENVHK-HSRLEAVNLQSGDILPQYKQEARKTPRHILHYCYFKTQWDMITILIRPY	230
QY	421 TAVFTPYSAFLKTEEGSPRATTEGGYACQPLAVVDLYDIMEFYDILINFRTTYNANE	480	QY	421 TAVFTPYSAFLKTEEGSPRATTEGGYACQPLAVVDLYDIMEFYDILINFRTTYNANE	480
DB	231 TAILVPYVVSFRTKNNVNA-----WLVDSDYDVFYDLVNLNFTFTVGPAG	278	DB	231 TAILVPYVVSFRTKNNVNA-----WLVDSDYDVFYDLVNLNFTFTVGPAG	278
QY	481 EVVSHPGRLAHVYFPGWFLIMVVAIPDGL-----IFSGSSELIGLKTATRLRIVRA	536	QY	481 EVVSHPGRLAHVYFPGWFLIMVVAIPDGL-----IFSGSSELIGLKTATRLRIVRA	536
DB	279 EVISDPKILIRMYLKTWFVIDLSCLPDYDVAFENVDBGISLTSLSLKAVRLRLGRVA	338	DB	279 EVISDPKILIRMYLKTWFVIDLSCLPDYDVAFENVDBGISLTSLSLKAVRLRLGRVA	338
QY	537 RKLDRSEYGAALVFLMLCTFALILAHMLACIYAIAGNMQPHMDSI-----GMLNHLGQ	592	QY	537 RKLDRSEYGAALVFLMLCTFALILAHMLACIYAIAGNMQPHMDSI-----GMLNHLGQ	592
DB	339 RKLDRHYTERGAALVLLVLCYGFGLAHNMAGCIYISGDEYIEPEDIKRTIRNNSMWLYQALAD	398	DB	339 RKLDRHYTERGAALVLLVLCYGFGLAHNMAGCIYISGDEYIEPEDIKRTIRNNSMWLYQALAD	398
QY	593 IGRPY--NSSGL-----GGPSIKDKYVTAALYFTSSLSYGVGNVSPNTNSEKIFISCVML	646	QY	593 IGRPY--NSSGL-----GGPSIKDKYVTAALYFTSSLSYGVGNVSPNTNSEKIFISCVML	646
DB	399 IGTPIQFQNGSGSGKMEGPPSKNSYVSISSLYFTMTSLTSYGFGNIAIPSTDIKIFAVAIM	458	DB	399 IGTPIQFQNGSGSGKMEGPPSKNSYVSISSLYFTMTSLTSYGFGNIAIPSTDIKIFAVAIM	458
QY	647 IGLSLMAYIFGVNSAIIQRLYSGTARYTQMLRAVEPLIRFHQIPNPLQRLREYFOHAMS	706	QY	647 IGLSLMAYIFGVNSAIIQRLYSGTARYTQMLRAVEPLIRFHQIPNPLQRLREYFOHAMS	706
DB	459 IGSLLYATIFGVNTTIFQOMYANTRYHEMLNSVDFKLXYOVPGLSERVADYIVSTWS	518	DB	459 IGSLLYATIFGVNTTIFQOMYANTRYHEMLNSVDFKLXYOVPGLSERVADYIVSTWS	518
QY	707 YTNIGDMAVLYKGFECQADICILNLSILOHCKPFGATGCGRALAMKRTKHAPP	766	QY	707 YTNIGDMAVLYKGFECQADICILNLSILOHCKPFGATGCGRALAMKRTKHAPP	766
DB	519 MSRGIDTEKVLQICPKDKMRADICVHLNKKVKEHNAFRLASGCRALAMEQYIHCAPG	578	DB	519 MSRGIDTEKVLQICPKDKMRADICVHLNKKVKEHNAFRLASGCRALAMEQYIHCAPG	578
QY	767 DTLVHAGDLTALFYISKGSIEILGSDVVVALIGNDFJGEBLNLVYARPKSGNSGVDRALY	826	QY	767 DTLVHAGDLTALFYISKGSIEILGSDVVVALIGNDFJGEBLNLVYARPKSGNSGVDRALY	826
DB	579 DLIYHAGSVDLSLCVVGSGSLEVIQDDVEVALILGQVGFYFKMEATLTAQSCAVNVALY	638	DB	579 DLIYHAGSVDLSLCVVGSGSLEVIQDDVEVALILGQVGFYFKMEATLTAQSCAVNVALY	638
QY	827 YCDLHKIHRDDLLEVLDMPPEFSDFHFWSSLEITNLR-----DTN-----	866	QY	827 YCDLHKIHRDDLLEVLDMPPEFSDFHFWSSLEITNLR-----DTN-----	866
DB	639 YCDLHVIRKDALOKLVEFYTAFSHSFNLLITFYMLRKRIYFRKISDYKREDEEREMKRN	698	DB	639 YCDLHVIRKDALOKLVEFYTAFSHSFNLLITFYMLRKRIYFRKISDYKREDEEREMKRN	698

```

QY 902 -----PGEVSALGPGRAGAGPSSR----- 920
Db 759 SVTVRESAPATPVFOASTSTVSDHAKLHAPGS -ECLGPKAGGADPAKKGNARFKDAC 817
QY 921 GRGGGGE-----SPSSGPPSSPESEDEGPGRRSSPLRLVPFSSDRPPGEP 967
Db 818 GK-GEOMNKVSKAESMETLPERTKRATKTKDSCDSCGIFTKSDRLDHWGCEARSPD- 875
QY 968 PGEPLMEDEKS 980
Db 876 --RSPILAEVYKHS 886

RESULT 9
148912
potassium channel subunit - mouse
C:Species: Mus musculus (house mouse)
C>Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 05-Nov-1999
C:Accession: 148912
R:Warme, J.W.; Ganetzky, B.
Proc. Natl. Acad. Sci. U.S.A. 91, 3438-3442, 1994
A>Title: A family of potassium channel genes related to eag in Drosophila and mammals.
A:Reference number: A54953; MUID:94211879; PMID:8159766
A:Accession: 148912
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-989 <RES>
A:Cross-references: EMBL:U04294; NID:9487739; PIDN:AAA62474.1; PID:9487740
C:Genetics:
A:Gene: m-eag

Query Match 23.1%; Score 1402; DB 2; Length 989;
Best Local Similarity 29.7%; Pred. No. 1.8e-80;
Matches 361; Conservative 166; Mismatches 363; Indels 322; Gaps 27;

QY 4 RGHVAPONTFLDTIRKFEQGRKFTIANARVENCAYICNDGFCELGYSRAEVNORP 63
Db 7 RRLVLAQONTFLDNVIR--SNDTNFVLAQAIDVMPVIYVSNDFCKSLGSHRAEVNQS 64
QY 64 CTGDFLHGPRTORRAAQIQAALLGAERKEVETAFYKDKSCFLCYDVVPVKNEDGAVI 123
Db 65 SASCFMAGELTDQYVEKVRQTFENYEMNSFEILMKYKKNRTPVWFYKIAPIRNEODKV 124
QY 124 MFLINFEVMEKDMVGSPPADTNHGRPPTSMLAGRAKTRFLKPLALALTARESSVRS 183
Db 125 LFLCTRS----- 131
QY 184 GAGGAGAGVAVVDVLTTPAAPSSESLALDEVTAMDNHVALGP-AEERRALVPGSPR 242
Db 132 -----DITAFKQPIE-----DDSCGKGWGFALTRAL----- 158
QY 243 SAGGQLPSPRAHSLNPASSSCSGLATRSRESCASVRRASADDEAMAGVLPPEPRH 302
Db 159 -----TSSR-----GVL----- 165
QY 303 ASTGAMHPLRSGLNSTSDSLVRYRTISKIPQTLINFDLKGDPFLASPTSDREITAP 362
Db 166 -----QQLAPS 171
QY 363 IK--ERTHNTEKVTQVLSGADVLPKYLAQPIRHMTLHSPKAVAWDMLLLVIY 420
Db 172 VQKGENVHK-HSRLEAVLQIGSDILPOYKQEARPTPHIILHCVFTWDMWILLITFY 230
QY 421 TAVFTPSAFLKETEGRPRATBCGACPLAVVDIIVIMFVDLILFRITTYVANE 480
Db 231 TALVLPVNSFKRONVA-----WLVPVSIVDVILFVLDIVLNFHTTFVGPAG 278
QY 481 EVVSHPRIVAHYFKGFLIDVAAIPEDL-----IFG----- 514
Db 279 EVISDRPLIMNLIKTFVIDLISCLPYDIINAFENVDEVSAFMGDPGKIGFADQIPPL 338
QY 515 -----SGSEBLIGLKTARLLRLVVARAKLDRYSEYGAVALFLMCTFALIAHMLACIMY 569

```

```

Db 339 EGRESOGISSLFSKLAVRLRLGRVARAKLDHYIEYGAAYLVLLYCVGLAAHMACIMY 398
QY 570 AIGMNEQPHMDSRI-----GWLHNLGDQIGPY--NSGL-----GGPSIKKUYVALYFTE 619
Db 399 SIGGYELFEDDTKIRNNMSWLYOLADLIGPYOPNGSGSGKWEGPSKNSVYISLFTM 458
QY 620 SLSLVGFGVNSPPTNSEKIFESICVMLIGSLMYSIGVNSAIIQRTGYSTARYHTQMLR 679
Db 459 TSLTSVFGNIAIPSTDLEKIFPAVALIMIGSLTYITIGNTYTFIQQYATNKRHEMLNS 518
QY 680 VREFIRPHQIPLNRQRLLEYFOHAWSYTNGIDMNAVLKGPPECLQADICLHLNRSLLQ 739
Db 519 VRDFLKLXYQPKGLSERVMYIVSTWMSRIGIDEKVIQJCPKMDRADICVHLNRKVEKE 578
QY 740 CKPRGATKGLRALAMKFTTHARPDDTLVHAGDLTALYFSRGSIELRGVVAAIL 799
Db 579 HPAERLASDCGLRLAMEFQTVHCAPDILYHAGEVSUSLCFVWSGLEVIQDEVAAIL 638
QY 800 GNDIFGEPLNLVAPKSGNDVRLATYCDLHKTHRDDLEVDIMPEFSDHFWSSLEIT 859
Db 639 GKGDVEGDVWKEKTLTAQSCANRALTYCDLHVKRALQKVEFTAFSHSPRNILIT 698
QY 860 FNLR-----DTN-----MIPSGSTLELGGFSRQRRKLSFR-- 893
Db 699 YNLKRLIVFRKISDVKREERMRKKNEARLILPPDHVRRLRFORQKEARLAAERGG 758
QY 894 --RTDKTEQGEVSALGPRAGAPSS-----RRPGGPMGESSSPSESEDEGPG 947
Db 759 RDLDLDVEKGNALTDHTSANHSLVKASVTVTRSPTPVPSFOAATYSTSDHAKLHAPG 818
QY 948 RSSSPLRLVPFSSPRPPEPPGCEPLMEDEKSSDTGNPLSGAFSGVSNIFSWGD-SRG 1006
Db 819 SEC-----LGRKAVSCDPAKKGMARFRDACKGKD-----WNKSKA 856
QY 1007 ROYQELP-KCPAPPSLLNIPLSSPGRRPGDVESRLDALQRLNRLRLSADMAVQL 1065
Db 857 ESMETLPERKAPCEATLKTKDSCDSCGIFTKSDL--RID-----NVGETRSPQORSPILA 908
QY 1066 LLOROMTLVPRAVSATVTPPGPRTSPPLVPSPPLTLTDSLVSQVQFMAECCLPPGAP 1125
Db 909 EVKHSFYPIEQTLQAVLVLEVYKELKEDIKALNAKMYSIEKQISEILLRIL--MSRGA 964
QY 1126 ELPOEGPTRRSLP 1139
Db 965 QSPQE--TGETSRP 976

RESULT 10
142394
potassium channel protein eag homolog - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999
C:Accession: T42394
R:Waterston, R.
submitted to the EMBL Data Library, June 1999
A:Reference number: Z22153
A:Accession: T42394
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-934 <NAT>
A:Cross-references: EMBL:AF036695; PIDN:AA88348.1
C:Genetics:
A:Introns: 27/1; 66/2; 104/1; 147/1; 164/3; 222/2; 263/3; 301/3; 410/1; 468/2; 667/3;
A>Note: F10B3.1

Query Match 23.0%; Score 1401; DB 2; Length 934;
Best Local Similarity 29.6%; Pred. No. 1.9e-80;
Matches 352; Conservative 167; Mismatches 304; Indels 366; Gaps 27;

QY 1 MPV-RGHVAPONTFLDTIRKFEQGRKFTIANARVENCAYICNDGFCELGYSRAEV 59
Db 1 MPVGRGLVAPONTFLDNVIRRCNNADTSEILANAAQVVDVPIYVNCNDGFSKLYGYTBAEI 60

```

QY	60	QORPCTDPLNGRPKIORRAAADIADALLGAERKYEIAFYRRDGCSEFLDVLVDVYPKNE	119
Dd	61	MOKPCLSAFMHEHEVEGVSLOMOALEARTQEAGIGLCKKNKTPITWLLVHAPITKNK	120
QY	120	GAVIDFIINFEVMEKDM--VGSPPADHTNRHQPPTSMILAPGRAKTFRLPLPALTARE	177
Dd	121	DAVLYLQGF-----KDITPLKQPLDDENK-----ALFCYTGRA	155
QY	178	SSVNSGGAGACAPGAVVVDVLDTPAADSSSESLADEVTAMDNHVAGLGPAERRALVGP	237
Dd	156	N-----AVEEYTRV-----	164
QY	238	GSPRSAPGOLPSPRAHSLNPDASGSSCSLARTRESRESCASVRRASADDIEMRAGVLP	297
Dd	165	-----	164
QY	238	PPPRHASTGAMHPLNSGLNSTSDLVKRYRTISKIPIQITLNFVDLKGDFLASPSDRE	357
Dd	165	-----	164
QY	358	IIAPKIKERTHNVTAKYQVYLSGLADVPEYKLOAPRIHRKMTILHSPKAWMDLILLL	417
Dd	165	-----VNLGDMPLPYOETPKTSPHILHSSKFTWDSITLAL	205
QY	418	VIYTAFTPYSAAFILKETEEGPAPTECGYACAPALVADLYIDIMFYILINFRTYYN	477
Dd	206	TEYTAFTMPFNTAF--KNSREMP-----CGGIDVALMDSIYDYFADILNFHTTFPG	259
QY	478	ANEVYSHPGRIAVHYEFKGMFLIDVAAIPEDLLIFGSGSEBLIG--LTKTARLRIV	533
Dd	260	PGGEVAVIPSYRQYFMSFLIDISCLPYDIFPMFKRDEIRISLFSALKVVRLLIRBG	319
QY	534	RVARKLDESEYGAVNLFMLCTFALIHMTACIYVAGNM--QJHMDRI--GWLJHL	589
Dd	320	RVARKLIDYLEGATLILLICAIYIVAHMLACVWFYIGDSEYRLKMDMLAPDGLMKL	379
QY	550	GDQIGKPPN-----SSGLGSPSIKDKYVLTALYFTSSITVSGFNVSPNTSEKIFISIC	643
Dd	380	SNDLROHNIPLSNKTTLVGGSPRSAYISSLYTMSGTSYGFGNISNTNEKIFGVC	439
QY	644	VMLGSLMYAIFGCVSAIIORLYSGTRYHQMRLVRPEFIHOJIPNLRQLEXPCH	703
Dd	440	MMIISALYIAIFGMMTIIIDOMTSYVRHEMISNREFILQDELPELARVMDYVVS	499
QY	704	AMSYTNGIDMANVALKGPEECLOADICLHNSRSLONCHKPEFGATGCLRALAMKFKTTA	763
Dd	500	TWAMTKGIDTAKVLGVCYPKDMKADICVHLNRKVEHNSCFRLASDGLRSLAMPLELNA	559
QY	764	PPGDIIVHAGDILLTALFYISGRSDEIILAGDVVALLGKNDITGEPL--NLTYAPGKSNQDY	822
Dd	560	APGDILYHTGESVDLMEVVSGSLEVIDDEDEVALLLGADVDEFGWANGSTGOSAAVY	619
QY	823	RALTYCDLHAKIHRDOLLVLDMYRPFSSFWMSLITETLRTNNIPGSPSTEEGGRS	882
Dd	620	RALTYSDLHMIAKKDMLMDVDEYKAFANSFANMMLTYNLTH--RAKFRVADYKKEELD	678
QY	883	KQRK-----RKLSPRR-----TKDTEOQGEVSALGGRAGA	915
Dd	679	AKRREKELTLPNDRHRIKLLFRMRHRGPRIIPSPMFADIEGKLKSTESIRI-----	731
QY	916	GPSSNGRGCGWGESPPSSPESBEDGGRSSPLRLVYFSSPRPPGEPGGEPLME	975
Dd	732	-----SSLHSMIDETGGGSSYVYK--SPRSKPKRP-----PLMK	763
QY	976	DCEKSDTCNPLSGAFSGVSNIFSEWG--DSRGROQELPRCAPPSLINLIPSGGRP	1033
Dd	764	RQYVEDALSRSTS-----WGDKDKDREMSLSLNIKTEKKSXFDI-----IGERL	807
QY	1035	R--GDVERLDALOR-----QJNRETRLSDMAIVYLOLL-----	1067
Dd	808	TIIOINSRLALLEVILGNNGGANTPSTMPVGSFSAINESGNSRITLTLAAPARVSWS	867

```

QY      1068 -----QR-----QMTLVPPAYS-----AATTCGPGCTSTSPLLP 1096
          ||| : ||| | | | | | | | |
Db      868 QHQPMPRFVEIYSETSTVPPLRELAEGEWEPPIREFTPNPSTSSRPV 916

RESULT 11
S52072
DmCNGC protein - fruit fly (Drosophila sp.)
C:Species: Drosophila sp.
C:Date: 14-Jul-1995 #sequence_revision 21-Jul-1995 #text_change 16-Jul-1999
C:Accession: S52072
R:Baumann, A.; Fritings, S.; Godde, M.; Selfert, R.; Kaupp, U.B.
EMBO J. 13, 5040-5050, 1994
A>Title: Primary structure and functional expression of a Drosophila cyclic nucleotid
A:Reference number: S52072; MUID:95045396; PMID:7957070
A:Accession: S52072
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-665 <BAU>
C:Genetics:
A:Gene: Flybase:Cng
A:Cross-references: Flybase:Fggn0014462
C:Superfamily: cyclic nucleotide-gated channel; cAMP receptor protein cyclic nucleoti
F:429-553/Domains: cAMP receptor protein cyclic nucleotide-binding domain homology <CA
Query Match       7.4%; Score 451.5; DB 2; Length 665;
Best Local Similarity 25.3%; Pred. NO. 1.1e-20;
Matches 147; Conservative 105; Mismatches 253; Indels 77; Gaps 20;

```

RESULT 12
A55251
cyclic nucleotide-gated Ca++ channel protein CNG-3 - bovine

C:Species: Bos primigenius taurus (cattle)
 C:Date: 06-Feb-1995 #sequence_revision 06-Feb-1995 #text_change 13-Aug-1999
 C:Accession: A55251; #S43976
 R:Biell, M.; Zong, X.; Distler, M.; Bosse, E.; Klugbauer, N.; Murakami, M.; Flockerzi, V.
 Proc. Natl. Acad. Sci. U.S.A. 91, 3505-3509, 1994
 A:Title: Another member of the cyclic nucleotide-gated channel family, expressed in test
 A:Reference number: A55251; MUID:94224768; PMID:8170936
 A:Accession: A55251
 A:Status: nucleic acid sequence not shown
 A:Molecule type: mRNA
 A:Residues: 1-706 <BIE>
 A:Cross-references: GB:X76485; NID:9488728; PIDN:CAA54023.1; PID:9488729
 A:Experimental source: kidney
 R:Weyand, I.; Godde, M.; Frings, S.; Weiner, J.; Mueller, F.; Altenhofen, W.; Hatt, H.;
 Nature 368, 859-863, 1994
 A:Title: Cloning and functional expression of a cyclic-nucleotide-gated channel from man
 A:Reference number: S43976; MUID:94211295; PMID:7512693
 A:Accession: S43976
 A:Status: preliminary; nucleic acid sequence not shown
 A:Molecule type: mRNA
 A:Residues: 1-706 <MEY>
 A:Cross-references: GB:X89600; NID:9908023; PIDN:CAA61759.1; PID:9908024
 A:Experimental source: testis
 C:Genetics:
 A:Gene: CNGB3
 C:Superfamily: cyclic nucleotide-gated channel; cAMP receptor protein cyclic nucleotide-
 C:Keywords: cAMP binding; ion channel; ion transport; membrane protein
 F:501-625/Domain: cAMP receptor protein cyclic nucleotide-binding domain homology <CAP>
 Query Match 7.28; Score 440; DB 2; Length 706;
 Best Local Similarity 23.3%; Pred. No. 6.3e-20;
 Matches 162; Conservative 115; Mismatches 283; Indels 134; Gaps 21;

```

QY 191 PGAVVVDL-----TPAASSESLADEVTADNVAAGPAREERALLVGPSP 241
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 16 PSYRTMDRDLDCLENGSLRHLPCETSSSELQEIAMETR---GLAESRQSSFTSGPT 71
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 242 RSAPGOLSPRAHS---LNDPASGSSCSLARTSRESCASVRRASADDEAMRAGVLP 298
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 72 RLISRLAMSARLHEDQRPDSFLERFG---AELQESSHSHVQFVGSGEP 126
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 299 PPRHASTGAMHPLRSGLNSTSDSLVRYRTISKIPQITLNFVDLKGDPPLASPTDREI 358
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 127 PDGGRSAMPLARNNTNCSSEKDKAKKEKEKE-----EKKENPKKEKKDSV 179
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 359 IAPRIKRTNVTETKYQVLSIGADVLPEYKLOAPRIHRWTLIHSFFKVMQ----- 411
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 180 MDSSSNMYHWLT-----VIAVPVYMWCL--VCRACFDELQSEHL 219
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 412 -WILLLVITYAVTFPYSAFLKETEGRPATECGYACOPLAVDLIVIMEIYVLI 469
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 220 MLMLVL-----DYSA-----DILYGMMLV 239
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 470 NFRITYNANEVYSHGRVAVHYFKG-WFLIDMVAALPPDLIFSGSEELGLKTAR 528
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 240 RARGLFEQG-LVMWASRLMKHYQTLFHLKDLVLSVPTDLVYFKGMN--YPELFPNR 296
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 529 LLRLVVARKLDRS-----EYGAVLFLMCTFALLAHMLACIWAIGMEQPHM 579
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 297 LKLIARLFEFFDRLETRTNPNNFRIGNLVLYLI-----IIHNNACIYFALSKIFGFT 351
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 580 DSRIGLWNLGDIKGRYNSGSGSPSIKDKYVYALYFTFSSLSVSGFVNSPTNSEKI 639
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 352 DS---WVYP-----NYSNPEYGLSKRYIYSLWSTLTLTIG-ETPPYKDEYL 398
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 640 FSTICWMLIGSLMTASIFGNVSAIIQRLYSTARYHTQMLVREFIRFHQIPNPLRQLEE 699
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 399 FVVIDFLVGLVIFATIGNGSMISNMNASRAEFQAKIDSIKQYMKRKYTKDLETRVIR 458
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 700 YFQHAMSTNGIDMNAVLKGFPECLADICLHNRSLQHCPRRGATKCLRLAMKFK 759
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 459 WEDYLMANKKTVEDEKVLKSLPKLKAELAINVHLDLTLRRVRLFDCEAGLIVELVILKR 518
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |

```

QY 760 TTHAPGDTLVHAGDLTLATLFFISRGSIELRGDYYV--AIIKKNIDFGEPLNTYARPG 817
 || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 DB 519 PAVSPEDYICKKKDDIGREMIIEGKLAVAEIGITQFVVLGSGISFGEISILNIKSG 578
 QY 818 S---NGDVALTYCDLHKIRHDDLLEVLDMYPE 847
 | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 DB 579 SGNRTANIRISIGYSDLFCLSKDDLMEALTEYPE 612
 || | | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 13

cyclic nucleotide-gated channel protein - human
 C:Species: Homo sapiens (man)
 C:Date: 14-Apr-1998 #sequence_revision 24-Apr-1998 #text_change 16-Jul-1999
 C:Accession: S74179
 R:Yu, W.P.; Grunwald, M.E.; Yau, K.W.
 FEBS Lett. 393, 211-215, 1996
 A:Title: Molecular cloning, functional expression and chromosomal localization of a h
 A:Reference number: S74179; MUID:96409310; PMID:8814292
 A:Accession: S74179
 A:Status: not compared with conceptual translation
 A:Molecule type: mRNA
 A:Residues: 1-695 <YUW>
 A:Experimental source: retina
 C:Genetics:
 A:Map position: 2
 C:Superfamily: cyclic nucleotide-gated channel; cAMP receptor protein cyclic nucleoti
 C:Keywords: ion channel; ion transport; membrane protein
 F:482-607/Domain: cAMP receptor protein cyclic nucleotide-binding domain homology <CA

Query Match 7.18; Score 433; DB 2; Length 695;
 Best Local Similarity 25.68; Pred. No. 1.7e-19;
 Matches 129; Conservative 103; Mismatches 204; Indels 68; Gaps 17;

```

QY 362 KIKERTNVTETKYQVLSIGADVL-PEYKLOAPRIHRWTLIHSFFKAVMDLILVLY 420
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 138 KCNTNTSNNTSEEEKTKKKKALIVDPSSNL---YFMLAIALP--VFYMWYLL---IC 188
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 421 TAVTFPYSAAPFLKETEGRPATECGYACOPLAVDLIVIMEIYVDTLIRRTYVANE 480
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 189 RACFDELQSEYLM-----WLVLDYSADLVLYLDVLRVRRGFLEQG- 230
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 481 EVVSHPRIVHY-FKGMFLDMVAALPPDL-LIFSGSEELIGLKTARLLRVAR 537
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 231 LMSVDTNRIMQHYTTQTFKLDVLSVPTDLATLKVGTNPE---VRFNLLKFSLFE 286
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 538 KLDRTS-----EYGAVLFLMCTFALLAHMLACIWAIGMEQPHMDSRIQMLHN 588
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 287 FFDRTETRTNYPNNFRIGNLVLYLI-----IIHNNACIYFALSKIFGFTDS--WVY- 337
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 589 LGDQIGKRYNSGSGSPSIRKDKYVYALYFTFSSLSVSGFVNSPTNSEKIFSTICWMLIG 648
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 338 --PWISIPENG-----RLSKRYIYSLWSTLTLTIG-ETPPYKDEYLIVYVDLVG 388
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 649 SLMTASIFGNVSAIIQRLYSTARYHTQMLVREFIRFHQIPNPLRQLEEYFQHAMSTY 708
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 389 VLIPIATIGNVNGSMISNMNASRAEFQAKIDSIKQYMKRKYTKDLETRVIRWPEYLMANK 448
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 709 NGIDMNAVLKGFPECLADICLHNRSLQHCPRRGATKCLRLAMKFKTTTHAPGDT 768
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 449 KTYDEKVELKSLPKLKAELAINVHLDLTLKVRIFQDCEAGLIVELVLRPFVFSFGDY 508
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 769 LVHAGDGLTLATLFFISRGSIELRGDYYV--AIIKKNIDFGEPLNTYARPGS---NGDY 822
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 509 ICKKGDIGKERYITINEKLVAVADDGVTQFVVLSDGSYFGEISITLNIKSGSKGNRRYANI 568
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 823 RALTYYCDLHKIRHDDLLEVLDMYP 846
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 569 RSIGYDLPFLSKDDLMEALTEYPE 592
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |

```

RESULT 14
 S55349

potassium channel protein KST1 - potato
 M:Alternate names: K+ channel inward rectifying protein KST1
 C:Species: Solanum tuberosum (potato)
 C:Date: 23-Aug-1995 #sequence_revision 19-Oct-1995 #text_change 21-Jul-2000
 C:Accession: S55349
 R:Mueller-Reber, B.; Ellenberg, J.; Provert, N.; Wallmitzer, L.; Busch, H.; Becker, D.;
 EMBO J. 14, 2409-2416, 1995
 A:Title: Cloning and electrophysiological analysis of KST1, an inward rectifying K(+) ch
 A:Reference number: S55349; MUID:95300773; PMID:7781596
 A:Accession: S55349
 A:Molecule type: mRNA
 A:Residues: 1-688 <MUD>
 A:Cross-references: EMBL:X79779; NID:9861146; PIDN:CAA56175.1; PID:9861147
 C:Genetics:
 A:Gene: KST1
 C:Keywords: potassium channel; potassium transport; transmembrane protein; transport pro

Query Match 7.0%; Score 428.5; DB 2; Length 688;
 Best Local Similarity 26.5%; Pred. No. 3,2e-19;
 Matches 134; Conservative 100; Mismatches 210; Indels 61; Gaps 16;

QY 365 ERTNHWTEKVTQVLSGADVLPYKIQAPRIHRTILHYSPEKAWDWLILLYITAVF 424
 | | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
 DB 23 ETTSGGFSSDLPLSLGARL--NY---ATKLRFTISPFNRYRCWEMELVMTYIAMI 77
 | | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
 QY 425 TPYSAFLIKETESGPPATEGCGYACQPLAVVDLYIDIMFYDILINFTTYNANE--VV 483
 | | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
 DB 78 SLFEVAFLSYRKDD-----TLFIYDNIYDCFFAIDILITFEVAYLHRSYLV 125
 | | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
 QY 484 SHPGIAVHYFKGWFLLDMVAIIP--DLLIFSGSEELIG---LLKTAARILRIYVARK 538
 | | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
 DB 126 DEPKIARIRYSTWFIYFCSTVPPQSILVFTGKESGCGVGRLLSLMKRLKRVSL 185
 | | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
 QY 539 LDRYSEYGAVALFLMCT-----FALIAHMLACIWAIGNMEOPHMDSRIGMLHNGD 591
 | | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
 DB 186 FARLEKDIRFMYFWTRCKLVSVTLFAV--HCAGCINMADR--YDSKKT----- 234
 | | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
 QY 592 QIGRPYNSGIGLP-----SIKDKYVTLVFTFSSLVGVGNVSPNNSEKIESICVWL 646
 | | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
 DB 235 -IGAVY-----PDFKOLSVGDRYITSLWSIVTLTTGYGDLHAENSREMLFDIFYML 286
 | | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
 QY 647 IGLMYASIFGNVSAIIOLKVSCTARYTOMLRVREFIRFHOIPNLRORLEEYQOHAMS 706
 | | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
 DB 287 ENLGLFTYITIGMTLVVHWTSPRTNREAVAKADEFKRNOLDPPRODVLSHMCLFK 346
 | | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
 QY 707 YTNGIDMNAVLKGEPELQADICLHLNRSLLQHCPEFGATKGCCLALAMKFTTHAPG 766
 | | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
 DB 347 -TEFLKQEEETLNGLPKAIKRTSIANHLFPPIVQNVHLFGVSRNLLFQVPEMEAEYPPK 405
 | | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
 QY 767 DTVHAGGLTALALFISGSIE--ILRGDVVAAILGK---NDIFGEPLMLYARPKSNGDV 822
 | | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
 DB 406 QEVILQNEAPPTDLYIVSGAVEFIQIIEGLOIIGKAVAGEIFGDIGVLCGRP--QPPAV 463
 | | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
 QY 823 RALFYCDLKHIRDDLEVLDMYPE 847
 | | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
 DB 464 RTTEISQILRLSRISLMIILANPE 488
 | | | | | : | | | | | : | | | | | : | | | | | : | | | | | :

RESULT 15

150630
 alpha subunit of cone photoreceptor CNV-channel - chicken
 C:Species: Gallus gallus (chicken)
 C:Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 13-Aug-1999
 C:Accession: 150630
 R:Bonigk, W.; Altenhofen, W.; Muller, F.; Dose, A.; Illing, M.; Molloy, R.S.; Kaupp, U.H
 Neuron 10, 865-877, 1993
 A:Title: Rod and cone photoreceptor cells express distinct genes for cGMP-gated channels
 A:Reference number: 150630; MUID:93264082; PMID:7684234
 A:Accession: 150630
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-735 <BON>
 A:Cross-references: EMBL:X89598; NID:9908850; PIDN:CAA61757.1; PID:9908851

C:Superfamily: cyclic nucleotide-gated channel; CAMP receptor protein cyclic nucleoti
 E:524-648/Domain: CAMP receptor protein cyclic nucleotide-binding domain homology <CA

Query Match 7.0%; Score 425.5; DB 2; Length 735;
 Best Local Similarity 27.1%; Pred. No. 5,5e-19;
 Matches 112; Conservative 93; Mismatches 163; Indels 45; Gaps 13;

QY 455 VDLIVDMFYDILINFTTYNANEVSHRGIAVHYFKG-WFLIDMVAIIPDL--L 511
 | | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
 DB 248 LDYCSDIYFDMFVRFRFTGFLEQG-LLVKDEKRLRDHYQTQVQFKLDVLSLLPPDLAVL 306
 | | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
 QY 512 IFGSGSEELIGLKTARILRLVVARKIDRYS-----EYGAVALFLMCTALALAH 562
 | | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
 DB 307 KLGLNYPE---LRPNRLRLARLEFPFDETRTNYPNMFRIQNLVYL----IIH 357
 | | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
 QY 563 WLACIWAIGNMEOPHMDSRIGMLH-NIG-DOIGRPYNSGIGSGSISKDKYVVALYFTFS 620
 | | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
 DB 358 WNACTIYPAISKYIGTGDS---WYIPNVSIPYGR-----LSRKTIYSILYWSVL 403
 | | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
 QY 621 SLTSVGFNVSPNTSEKIFISICVMLIGSLMAYASIFGNVSAIIQRLYSGTARYHTOMLRV 680
 | | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
 DB 404 TLTITIG-ETPPVKKDEYLFVVIDFLVGLIFATIVGNVSMNSRAEFQAKVDISI 462
 | | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
 QY 681 REPIFHQIPNPLRGLREYQOHANSTYNGIDMNAVLKGFPECLQADICHLNRSLLQHC 740
 | | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
 DB 463 KOYMHFRKVTYDLEARVAKMEDYLMTNKKYDEKEVLNLPDKLAEIAINVHLDTLKV 522
 | | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
 QY 741 KPEFGATKGCCLALAMKFTTHAPGDTLVHAGDLLTALYFISRSISIELRGDVV--AI 798
 | | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
 DB 523 RIFQDCEAGLLLELVALKLKPFTYFSPGDYICKKGDIKREMYITKEGKLAVVADDDGITQFVV 582
 | | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
 QY 799 LGRNDIFGEPLMLYARPKS---NGDYRALTYCDLKHIRDDLEVLDMYPE 847
 | | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
 DB 583 LSDGSYFGEISILINIKSGSKGNRRPTANIRISIGSDFLCSKDDLMALTEYPE 635
 | | | | | : | | | | | : | | | | | : | | | | | : | | | | | :

Search completed: September 23, 2003, 17:40:17
 Job time : 51 secs

Query Match	Best Local Similarity	Score	DB	Length	Mismatches	Indels	Gaps
100.0%	100.0%	6079	20	1159	0	0	0
Matches 1159; Conservative 0; Mismatches 0; Indels 0; Gaps 0							
Sequence	1159 AA:						
MPVRGHAPONTFLDTIIRKEPESGRKFIINARVENAVLYCNDGFCGLGYSRAEYM	60						
1 MPVRGHAPONTFLDTIIRKEPESGRKFIINARVENAVLYCNDGFCGLGYSRAEYM	60						
QRPCTCDFLHGPRTRQRAAQAQALALGAEEKVEIAFYRKDGSFCFLVDVVPVKNEDG	120						
61 QRPCTCDFLHGPRTRQRAAQAQALALGAEEKVEIAFYRKDGSFCFLVDVVPVKNEDG	120						
QRPCTCDFLHGPRTRQRAAQAQALALGAEEKVEIAFYRKDGSFCFLVDVVPVKNEDG	120						
AVIMFLNFEVYMEKDMVGSAPHDNNHNGPRTSWLAPGRKFRKLPLALLLTARESSV	180						
121 AVIMFLNFEVYMEKDMVGSAPHDNNHNGPRTSWLAPGRKFRKLPLALLLTARESSV	180						
AVIMFLNFEVYMEKDMVGSAPHDNNHNGPRTSWLAPGRKFRKLPLALLLTARESSV	180						
RSGGAGGAGAGAVVVDVLTTPAAPSSTSLADEVTAMDNHVAAGLGPAREALVGPSP	240						
181 RSGGAGGAGAGAVVVDVLTTPAAPSSTSLADEVTAMDNHVAAGLGPAREALVGPSP	240						
RSGGAGGAGAGAVVVDVLTTPAAPSSTSLADEVTAMDNHVAAGLGPAREALVGPSP	240						
PRAPGOLPSPRAHSLNDPASGSSCLARTSRESCASVRRASSADIDEAMRAGVLPSP	300						
241 PRAPGOLPSPRAHSLNDPASGSSCLARTSRESCASVRRASSADIDEAMRAGVLPSP	300						
PRAPGOLPSPRAHSLNDPASGSSCLARTSRESCASVRRASSADIDEAMRAGVLPSP	300						
RHASTGAMHPLRSGLNSTSDSLRYRTISKIPQITLNEVDLKDPLFASPTSPRETTA	360						
301 RHASTGAMHPLRSGLNSTSDSLRYRTISKIPQITLNEVDLKDPLFASPTSPRETTA	360						
RHASTGAMHPLRSGLNSTSDSLRYRTISKIPQITLNEVDLKDPLFASPTSPRETTA	360						
PKIKERTHNTEKYQVLSLGVADVLYPEKKLAPRIHRWTIILYSPFKAVMDLILLIYV	420						
361 PKIKERTHNTEKYQVLSLGVADVLYPEKKLAPRIHRWTIILYSPFKAVMDLILLIYV	420						
PKIKERTHNTEKYQVLSLGVADVLYPEKKLAPRIHRWTIILYSPFKAVMDLILLIYV	420						
TAFTTPYSAALFLKETEGSPRATCEGACOPPLAVVDLIVDIMEIVDILINFTTYVANE	480						
421 TAFTTPYSAALFLKETEGSPRATCEGACOPPLAVVDLIVDIMEIVDILINFTTYVANE	480						
TAFTTPYSAALFLKETEGSPRATCEGACOPPLAVVDLIVDIMEIVDILINFTTYVANE	480						
EVSHSPGRIAHYEFGWFLIDMVAALIPDDLIFGSGSEELGLTARLLRIVRARKUD	540						
481 EVSHSPGRIAHYEFGWFLIDMVAALIPDDLIFGSGSEELGLTARLLRIVRARKUD	540						
EVSHSPGRIAHYEFGWFLIDMVAALIPDDLIFGSGSEELGLTARLLRIVRARKUD	540						
RYSEYGAALVFLMCTFEALIAHWLACIUYAIGNMEQPHMDSRIGWLNHLDQIGRPYNS	600						
541 RYSEYGAALVFLMCTFEALIAHWLACIUYAIGNMEQPHMDSRIGWLNHLDQIGRPYNS	600						
RYSEYGAALVFLMCTFEALIAHWLACIUYAIGNMEQPHMDSRIGWLNHLDQIGRPYNS	600						
GLGSPSLKDKVYVTLVTFSSLSVGFQNSVPNTNSEKIFSTICVWLIGSLMAYASTFGVNS	660						
601 GLGSPSLKDKVYVTLVTFSSLSVGFQNSVPNTNSEKIFSTICVWLIGSLMAYASTFGVNS	660						
GLGSPSLKDKVYVTLVTFSSLSVGFQNSVPNTNSEKIFSTICVWLIGSLMAYASTFGVNS	660						
AIQRLKVSQTAHYHOMLRVEEFIEHQIPDLRQRLREYEQHNASYNGIDIMNAVLFGE	720						

Db		661	AIIQRLVSGTARYHTQMIRVAEFLIRFQIIPNPLRQRLEEFYQHAWSTNGIDMAVLKGF	720
Qy		721	PECLQADICLHLNSLLQHCPRFGATKGCLRALAMFKTTTHAPGDTLVHAGDLTALY	780
Db		721	PECLQADICLHLNSLLQHCPRFGATKGCLRALAMFKTTTHAPGDTLVHAGDLTALY	780
Qy		781	FISGSIETILKGDVVAVILGNKDIFGEPLNTIYARPGSNQGVRAITYCDLHKIHRDPLE	840
Db		781	FISGSIETILKGDVVAVILGNKDIFGEPLNTIYARPGSNQGVRAITYCDLHKIHRDPLE	840
Qy		841	VLDWYPESDHFWMSLETTFENLRDTNMIPGSPGSTEEGGFSRQKRKLSEFRRTDKDE	900
Db		841	VLDWYPESDHFWMSLETTFENLRDTNMIPGSPGSTEEGGFSRQKRKLSEFRRTDKDE	900
Qy		901	QPGEYSALGRPRAAGRSRRCPGPMGESPSGSSPESEDDEGPGRSSPLRLVPFS	960
Db		901	QPGEYSALGRPRAAGRSRRCPGPMGESPSGSSPESEDDEGPGRSSPLRLVPFS	960
Qy		961	PRPGEPPGGEPMLNEDCKSSDTCNPILSGAFSGSNIFFSMGDSRGROYELPCPAPTP	1020
Db		961	PRPGEPPGGEPMLNEDCKSSDTCNPILSGAFSGSNIFFSMGDSRGROYELPCPAPTP	1020
Qy		1021	SLLNPIUSSPGRRPRGDEVSRDALQROLNFLETRLSDMATVYLQLLRQMTLVPPAYSA	1080
Db		1021	SLLNPIUSSPGRRPRGDEVSRDALQROLNFLETRLSDMATVYLQLLRQMTLVPPAYSA	1080
Qy		1081	VTTGGPGGTSTPLPYSPPLTLRLDISQVSQPMACEBELPGAPBELPOBGPTRLSLP	1140
Db		1081	VTTGGPGGTSTPLPYSPPLTLRLDISQVSQPMACEBELPGAPBELPOBGPTRLSLP	1140
Qy		1141	QLGALTQSPLRHGSDPCS	1159
Db		1141	QLGALTQSPLRHGSDPCS	1159
<hr/>				
RESULT 2				
AAy85405				
ID	AAy85405	standard; Protein;	1159 AA.	
XX				
AC	AAy85405;			
XX				
DT	20-JUN-2000	(first entry)		
XX				
DE	Long QT syndrome associated HERG protein.			
KM	HERG; mutation; long QT syndrome; LQT syndrome; gene therapy; human.			
XX				
OS	Homo sapiens.			
XX				
FH	Key	Location/Qualifiers		
FT	Misc-difference	572 /note= "can be substituted with Cys; specifically claimed"		
FT	Misc-difference	588 /note= "can be substituted with Asp; specifically claimed"		
FT	Misc-difference	614 /note= "can be substituted with Val; specifically claimed"		
FT	Misc-difference	630 /note= "can be substituted with Ala; specifically claimed"		
FT				
FN	WO200006772-A1.			
PD	10-FEB-2000.			
XX				
PE	20-JUL-1999;	99WO-US16337.		
XX				
PR	27-JUL-1998;	98US-0122847.		
XX				
PR	06-JAN-1999;	99US-0226012.		
XX				

PA (UTAH) UNIV UTAH RES FOUND.
 XX Keating MT, Splawski I;
 XX WPI, 2000-195319/17.
 DR N-PSDB; AAA07601.
 XX
 PT New isolated mutant HERG nucleic acids, useful for developing products
 PT for the diagnosis, prevention and treatment of long QT syndrome
 XX
 PS Claim 11; Fig 8A-B; 163pp; English.
 XX
 CC The invention relates to a HERG protein having a mutation compared to
 CC wild-type HERG, and is useful for developing products for the diagnosis,
 CC prevention and treatment of long QT (LQT) syndrome. The products and
 CC methods can be used for the diagnosis of subjects with LQT syndrome. They
 CC can also be used to screen for drugs for treating or preventing LQT
 CC syndrome. The HERG nucleic acids can also be used for gene therapy and
 CC HERG peptides can be used for peptide therapy. The present sequence
 CC represents the LQT syndrome associated HERG protein. The HERG protein
 CC gene sequence comprising any of the following mutations apart from those
 CC indicated above is also specifically claimed for in the specification.
 CC The mutations arise from specific alterations in the encoding HERG gene
 CC sequence. The mutations can be F29L, N31T, C44S, G47V, G53R, R36Q,
 CC C66G, H70R, P72Q, R73frameshift, A78P, A83frameshift, Q81stop, L86R,
 CC P141frameshift, P151frameshift, P241frameshift, V295frameshift, R312C,
 CC P347S, R531Q, L552S, A561T, G584S, W585C, I593T, G604S, D609N, T613M,
 CC L615V, G626S, P632L, P632S, K638R, delK638, M645L, E682stop,
 CC H739frameshift, R752W, I798frameshift, F805S, F805C, R823W, N861I,
 CC K866frameshift, P917L, R920frameshift, R922W, G925frameshift,
 CC P968frameshift, P986frameshift, W1001stop, R1014stop, G1031frameshift,
 CC and P1101frameshift.
 CC
 XX
 SQ Sequence 1159 AA:
 Query Match 100.0%; Score 6079; DB 21; Length 1159;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1159; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 481 EVVSHPRIAVHYFKGFLIDMVAIPFDLLIFGSGSEELIGLTKTARLLRVAVARKLD 540
 QY 541 RYSEYGAAYFLIMCTFALLAHMLACIMYALGNNEQPHMDSRGMLNLDGQIGKPNSS 600
 DB 541 RYSEYGAAYFLIMCTFALLAHMLACIMYALGNNEQPHMDSRGMLNLDGQIGKPNSS 600
 QY 601 GLGSPSIKDKYVVALYTFESSLTSVCGNVPNNNSKIFISICVMILGSLMYASIFGNVS 660
 DB 601 GLGSPSIKDKYVVALYTFESSLTSVCGNVPNNNSKIFISICVMILGSLMYASIFGNVS 660
 QY 661 AIIQRLYSGTARYHTOMLRVREFTRHQIPNPLRQRLEEVYQHAMSTYNGIDMAVLKGF 720
 DB 661 AIIQRLYSGTARYHTOMLRVREFTRHQIPNPLRQRLEEVYQHAMSTYNGIDMAVLKGF 720
 QY 721 PECLQADICHLNLSLQHQCKPFGAKGCLRALAMFKTTHAPPGDTLVHAGDLTALY 780
 DB 721 PECLQADICHLNLSLQHQCKPFGAKGCLRALAMFKTTHAPPGDTLVHAGDLTALY 780
 QY 781 FISRGSIEILRGDVVAIILGKNDIFGEPMLLYARPGKSNGBVRLATYCDLHKIHRDDLE 840
 DB 781 FISRGSIEILRGDVVAIILGKNDIFGEPMLLYARPGKSNGBVRLATYCDLHKIHRDDLE 840
 QY 841 VLDMPPEPSDHFWSLEITFNLKDTNMIPOSPGSTLELGGFRORRKLFRRTDYTE 900
 DB 841 VLDMPPEPSDHFWSLEITFNLKDTNMIPOSPGSTLELGGFRORRKLFRRTDYTE 900
 QY 901 QPGVSAALGPRGAGSSNGRPGPGWGESPPSGSPSESEDEGPRSSSPLLVPPSS 960
 DB 901 QPGVSAALGPRGAGSSNGRPGPGWGESPPSGSPSESEDEGPRSSSPLLVPPSS 960
 QY 961 PRPEGPGEPEPLMEDEKESDTCNPLSGAFSGVSNIFSWGDSRGQOYELPRCPAPTP 1020
 DB 961 PRPEGPGEPEPLMEDEKESDTCNPLSGAFSGVSNIFSWGDSRGQOYELPRCPAPTP 1020
 QY 1021 SLNIPILSSPGRRRGVESRLDALQRLNLETRLRSADMAVYQLLQRLQMTLVPPAYSA 1080
 DB 1021 SLNIPILSSPGRRRGVESRLDALQRLNLETRLRSADMAVYQLLQRLQMTLVPPAYSA 1080
 QY 1081 VTPPGPGTSPPLPVSPPLTLTLDLSOVQPMACEELPGCAPELPOBGPTRRLSLPG 1140
 DB 1081 VTPPGPGTSPPLPVSPPLTLTLDLSOVQPMACEELPGCAPELPOBGPTRRLSLPG 1140
 QY 1141 QLGALTSQPLRHGSDPGS 1159
 DB 1141 QLGALTSQPLRHGSDPGS 1159

RESULT 3
 AAM51812
 ID AAM51812 standard; Protein: 1159 AA.
 XX
 AC AAM51812;
 XX
 DT 22-JAN-2002 (first entry)
 XX
 DE Human potassium channel subunit HERG.
 XX
 KW S. cerevisiae; yeast; potassium translocation system; TRK1; TRK2;
 KW TOK1; human erg potassium ion channel; HERG; antiarrhythmic;
 KW antiinflammatory; antidiarrilatory; HERG modulator; human.
 OS Homo sapiens.
 XX
 PN DEL9953478-AI.
 PD 11-OCT-2001.
 XX
 PF 06-NOV-1999; 99DE-1053478.
 XX
 PR 06-NOV-1999; 99DE-1053478.
 XX
 PA (LICH/) LICHTENBERG-FRATTE H.
 PA (LUDW/) LUDWIG J.

```

XX Lichtenberg-Frate H, Ludwig J;
PI
XX WPI: 2001-603577/69.
DR N-PSDB: AAI66257.
XX
XX Genetically modified yeast lacking endogenous potassium transport
PT activity, useful for identifying e.g. antiarrhythmic agents, includes a
PT functional human potassium channel
XX
PS Example 3: Fig 8: 40pp; German.
XX
CC The present invention relates to a genetically modified Saccharomyces
CC cerevisiae in which the endogenous potassium-translocation systems (TRK1,
CC TRK2 and TOK1) are specifically deleted and the human erg potassium ion
CC channel (HERG) is stably integrated and expressed. This can be used to
CC identify specific modulators of HERG, which are potentially useful as
CC antiarrhythmic, antifibrillatory and antiinflammatory agents. The present
CC sequence is the human HERG protein.
XX
SQ Sequence 1159 AA;
Query Match 100.0%; Score 6079; DB 22; Length 1159;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1159; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MPVRGHVAPQNTFLDTIRKFEQGRKFIITANAVENCAYICNDGFCELCGYSRAEYM 60
DB 1 MPVRGHVAPQNTFLDTIRKFEQGRKFIITANAVENCAYICNDGFCELCGYSRAEYM 60
QY 61 QRPCTCDLHGPRTRORRAAOIQAALLGAERKVEIAFYRKDSCFLCIVDVPVKNEG 120
DB 61 QRPCTCDLHGPRTRORRAAOIQAALLGAERKVEIAFYRKDSCFLCIVDVPVKNEG 120
QY 61 QRPCTCDLHGPRTRORRAAOIQAALLGAERKVEIAFYRKDSCFLCIVDVPVKNEG 120
DB 61 QRPCTCDLHGPRTRORRAAOIQAALLGAERKVEIAFYRKDSCFLCIVDVPVKNEG 120
QY 121 AVIMFILMEVYMEKDMGSPADHNTNHRGPTSMIAPGAKTFRLLKLPALLATARESSV 180
DB 121 AVIMFILMEVYMEKDMGSPADHNTNHRGPTSMIAPGAKTFRLLKLPALLATARESSV 180
QY 121 AVIMFILMEVYMEKDMGSPADHNTNHRGPTSMIAPGAKTFRLLKLPALLATARESSV 180
DB 121 AVIMFILMEVYMEKDMGSPADHNTNHRGPTSMIAPGAKTFRLLKLPALLATARESSV 180
QY 181 RSGGAGAGAGAVVVDVLTTPAAPSSESLADEVTAMDNHVAGLPAERRALVGPSP 240
DB 181 RSGGAGAGAGAVVVDVLTTPAAPSSESLADEVTAMDNHVAGLPAERRALVGPSP 240
QY 181 RSGGAGAGAGAVVVDVLTTPAAPSSESLADEVTAMDNHVAGLPAERRALVGPSP 240
DB 181 RSGGAGAGAGAVVVDVLTTPAAPSSESLADEVTAMDNHVAGLPAERRALVGPSP 240
QY 241 PRSAGQLPSPRAHSLNDPASGSSCLARTSRSCASVRRASADIDEAMRAGVLP 300
DB 241 PRSAGQLPSPRAHSLNDPASGSSCLARTSRSCASVRRASADIDEAMRAGVLP 300
QY 301 RHASGAMHPLRSGGLNSTSDVRYRTISKIPQITLNFVDLKGDPFLASPTSDREIT 360
DB 301 RHASGAMHPLRSGGLNSTSDVRYRTISKIPQITLNFVDLKGDPFLASPTSDREIT 360
QY 301 RHASGAMHPLRSGGLNSTSDVRYRTISKIPQITLNFVDLKGDPFLASPTSDREIT 360
DB 301 RHASGAMHPLRSGGLNSTSDVRYRTISKIPQITLNFVDLKGDPFLASPTSDREIT 360
QY 361 PKIKERTHNTEKTVQVLSLGDVLPYKLAQPRITHRTILHSPFKAVMDLILLVIY 420
DB 361 PKIKERTHNTEKTVQVLSLGDVLPYKLAQPRITHRTILHSPFKAVMDLILLVIY 420
QY 421 TAVFTPYSAALAKTEEGPRATEGACQPLAVVDLIVDIMEIVDILINERTTYVANE 480
DB 421 TAVFTPYSAALAKTEEGPRATEGACQPLAVVDLIVDIMEIVDILINERTTYVANE 480
QY 421 TAVFTPYSAALAKTEEGPRATEGACQPLAVVDLIVDIMEIVDILINERTTYVANE 480
DB 421 TAVFTPYSAALAKTEEGPRATEGACQPLAVVDLIVDIMEIVDILINERTTYVANE 480
QY 481 EVVSHPGIAVHYFGWFLIMVAAIPFDLLIFSGSSELLGLLTARLRIVRARAKLD 540
DB 481 EVVSHPGIAVHYFGWFLIMVAAIPFDLLIFSGSSELLGLLTARLRIVRARAKLD 540
QY 481 EVVSHPGIAVHYFGWFLIMVAAIPFDLLIFSGSSELLGLLTARLRIVRARAKLD 540
DB 481 EVVSHPGIAVHYFGWFLIMVAAIPFDLLIFSGSSELLGLLTARLRIVRARAKLD 540
QY 541 RYSEYGAVALFLMCTFALIAHWLACIWIYAIGNMOPHMDSRIGWLHNLGDOIGRPYSS 600
DB 541 RYSEYGAVALFLMCTFALIAHWLACIWIYAIGNMOPHMDSRIGWLHNLGDOIGRPYSS 600
QY 541 RYSEYGAVALFLMCTFALIAHWLACIWIYAIGNMOPHMDSRIGWLHNLGDOIGRPYSS 600
DB 541 RYSEYGAVALFLMCTFALIAHWLACIWIYAIGNMOPHMDSRIGWLHNLGDOIGRPYSS 600
QY 601 GLGPGSIXDKYVITALYFTFSSLTSGVGNVSPNTNSEKIFISICVWLIGSLMAYSTFGVNS 660
DB 601 GLGPGSIXDKYVITALYFTFSSLTSGVGNVSPNTNSEKIFISICVWLIGSLMAYSTFGVNS 660
QY 601 GLGPGSIXDKYVITALYFTFSSLTSGVGNVSPNTNSEKIFISICVWLIGSLMAYSTFGVNS 660
DB 601 GLGPGSIXDKYVITALYFTFSSLTSGVGNVSPNTNSEKIFISICVWLIGSLMAYSTFGVNS 660
QY 661 AIIORLYSGTARYHTOMLRVREFTRFHOIIPNPLRQRLREYQOHAMSYNGIDMNAVILKGF 720
DB 661 AIIORLYSGTARYHTOMLRVREFTRFHOIIPNPLRQRLREYQOHAMSYNGIDMNAVILKGF 720

```

```

QY 721 PECTIADICLHLNRSLLQHCPRGATKCCLRALAMKERTTHAPPGDTLVHAGDLLTALY 780
DB 721 PECTIADICLHLNRSLLQHCPRGATKCCLRALAMKERTTHAPPGDTLVHAGDLLTALY 780
QY 781 FISRGSIELKRDVYVALLGKNDIGCEPLNLYARPKSGNSGVRYALTYCDLKHHRDLE 840
DB 781 FISRGSIELKRDVYVALLGKNDIGCEPLNLYARPKSGNSGVRYALTYCDLKHHRDLE 840
QY 841 VLDMPERSDHFWSSLEITFNLRDTNMIIPGSPSTELGSGFSRQKRRLSPFRRTDKDTE 900
DB 841 VLDMPERSDHFWSSLEITFNLRDTNMIIPGSPSTELGSGFSRQKRRLSPFRRTDKDTE 900
QY 901 QPGEYSALGPRAGAGPSSRGRPGGPMWESPSSGSPSSSEDEGPGRSSSPLRLVPSS 960
DB 901 QPGEYSALGPRAGAGPSSRGRPGGPMWESPSSGSPSSSEDEGPGRSSSPLRLVPSS 960
QY 961 PRPGEPPGGEPLMDCERKSSDTCNPLSGATSGVSNITFSFWGDSRGROYOELPRCAPTP 1020
DB 961 PRPGEPPGGEPLMDCERKSSDTCNPLSGATSGVSNITFSFWGDSRGROYOELPRCAPTP 1020
QY 1021 SLNITPLSSPGRPRRGDVESRLDALORQNLRETRLSADMATVQLQLOMTLVPPAYSA 1080
DB 1021 SLNITPLSSPGRPRRGDVESRLDALORQNLRETRLSADMATVQLQLOMTLVPPAYSA 1080
QY 1081 VTPPGPPTSTSPPLPVSPPLTLTDSLSQVSOFNACBELPPGABELPOEGPTRTLSPG 1140
DB 1081 VTPPGPPTSTSPPLPVSPPLTLTDSLSQVSOFNACBELPPGABELPOEGPTRTLSPG 1140
QY 1141 QLGALTSQPLHRHSGDPGS 1159
DB 1141 QLGALTSQPLHRHSGDPGS 1159
RESULT 4
AAB70939
ID AAB70939 standard; Protein: 1159 AA.
XX
AC AAB70939;
XX
XX 22-AUG-2001 (first entry)
XX
DE Human eag-related protein HERG.
XX
XX Potassium transporter; K+ transporter; TKHp; mutant; potassium uptake;
XX selectable marker; auxotrophic marker; resistance marker; Trk2p;
XX eag related gene; HERG; potassium ion channel; antiarrhythmic;
XX antifibrillatory; antiinflammatory.
XX
OS Homo sapiens.
XX
XX DEL9941768-A1.
XX
XX 15-MAR-2001.
XX
XX 02-SEP-1999; 99DE-1041768.
XX
XX 02-SEP-1999; 99DE-1041768.
XX
XX (LICH/) LICHTENBERG-FRATE H.
XX
XX Lichtenberg-Frate H;
XX
XX WPI: 2001-246033/26.
XX
XX N-PSDB: AAF61965.
XX
XX New Schizosaccharomyces pombe mutants with defective potassium uptake
XX useful for expressing human eag related gene in screening assays for
XX potassium ion channel modulators -
XX
XX Example 1.4; Page 13; 24pp; German.
XX
XX This invention describes novel Schizosaccharomyces pombe mutants with
CC

```

defective potassium uptake which are obtainable by introducing one or more selectable markers (auxotrophy and/or resistance markers) into the genes for potassium transporter TRKp and/or Trk2p. The invention also describes (1) a genetically modified *S. pombe* strain that expresses human eag related gene (HERG) but does not express endogenous TRKp or Trk2p, (2) a screening assay for specific modulators of HERG potassium ion channel, comprising growing an *S. pombe* strain of type (1) in the presence and absence of a test substance and measuring any change in potassium transport. The mutants can be transformed to express human eag related gene (HERG) and used in screening assays for specific modulators of HERG potassium ion channel, which are potentially useful as antiarrhythmic, antifibrillatory and antiinflammatory agents. This sequence represents the human HERG protein described in the disclosure of the invention.

Sequence 1159 AA;

Query Match	100.0%;	Score 6079;	DB 22;	Length 1159;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 1159; Conservative	0;	Mismatches	0;	Gaps 0;

QY	1	MPVRRGHVAPQNTFJDTIIIRKFEQGSRKFIIANARVENCAVITYCNDGFCELCGYSRAEVM	60
Db	1	MPVRRGHVAPQNTFJDTIIIRKFEQGSRKFIIANARVENCAVITYCNDGFCELCGYSRAEVM	60
QY	61	QRPCTCDLHGRRTORRAAQAQIQAQALLGAEKKEVETAFRKQGSCELCIVDVVPAKNEG	120
Db	61	QRPCTCDLHGRRTORRAAQAQIQAQALLGAEKKEVETAFRKQGSCELCIVDVVPAKNEG	120
QY	121	AVIMEILNFEVYMEDMVGSPAHDPNNHRKPPISWILAPGRKTFRLKPLALALTARESSV	180
Db	121	AVIMEILNFEVYMEDMVGSPAHDPNNHRKPPISWILAPGRKTFRLKPLALALTARESSV	180
QY	181	RSGGAGGAGAPCAVAVVDVLTTPAAPSSESLADEVTAMDNHVAGLCPAEBERRALVGPSP	240
Db	181	RSGGAGGAGAPCAVAVVDVLTTPAAPSSESLADEVTAMDNHVAGLCPAEBERRALVGPSP	240
QY	241	PRSAAGOLPSPPAHSLNPDASGSCSILARTRESCASVRRASSADDITAMRAGVLPSP	300
Db	241	PRSAAGOLPSPPAHSLNPDASGSCSILARTRESCASVRRASSADDITAMRAGVLPSP	300
QY	301	RHASTGAMHPLNSGLNSTDSDLVRYRTISKIPIOTILNFVULKGPFLASTPSDBREIIA	360
Db	301	RHASTGAMHPLNSGLNSTDSDLVRYRTISKIPIOTILNFVULKGPFLASTPSDBREIIA	360
QY	361	PKIKERTHNVTBKVTQVITSLGADVLPBKLOAPRIHRWTLIHYSPKAVMDWILLVLIVY	420
Db	361	PKIKERTHNVTBKVTQVITSLGADVLPBKLOAPRIHRWTLIHYSPKAVMDWILLVLIVY	420
QY	421	TAVFPPYSAFLTKTEBEGRPATCEGYNCOPLAVVDLIVDIMFYDILLNFRITYYNAME	480
Db	421	TAVFPPYSAFLTKTEBEGRPATCEGYNCOPLAVVDLIVDIMFYDILLNFRITYYNAME	480
QY	481	EYVSHPRGIAVHYEFKGWFLIDMVAAIPEDLILFSGSSEELIGLTKTARLLIYRVARKID	540
Db	481	EYVSHPRGIAVHYEFKGWFLIDMVAAIPEDLILFSGSSEELIGLTKTARLLIYRVARKID	540
QY	541	RYSEYGAADVFLIMKCTFALLAHMLACIWAIAINMEOPHMDSIGWLHNIGDOIGKPYNSS	600
Db	541	RYSEYGAADVFLIMKCTFALLAHMLACIWAIAINMEOPHMDSIGWLHNIGDOIGKPYNSS	600
QY	601	GLGGESIKDKYVTALYTFSSLTSGFGVAVSNTSEKFEFCVMIIGSLMASTIFGNSS	660
Db	601	GLGGESIKDKYVTALYTFSSLTSGFGVAVSNTSEKFEFCVMIIGSLMASTIFGNSS	660
QY	661	AIQIKVYSGTAYVHTOMLRVRETFEIPHOQIPNBLRQLEEYEQHAMSXYTGIDMNAVLKGF	720
Db	661	AIQIKVYSGTAYVHTOMLRVRETFEIPHOQIPNBLRQLEEYEQHAMSXYTGIDMNAVLKGF	720
QY	721	PECLQADITCLHUNBSLLOHCKPFRGATKGCPLALAMKEFTTAPPEGDILVHAGDILLIY	780
Db	721	PECLQADITCLHUNBSLLOHCKPFRGATKGCPLALAMKEFTTAPPEGDILVHAGDILLIY	780

QY	781	FISGSGIEILRGDVVAALIGKNDIFEGEPNLNYPKGSNDGVYALTYCDLHKHTRDDLE	840
Db	781	FISGSGIEILRGDVVAALIGKNDIFEGEPNLNYPKGSNDGVYALTYCDLHKHTRDDLE	840
QY	841	VLDWYPEPSDHFWSLLETFNLRTNNI PGSPGSTELEGGFSRQRKKLSFRRTTDKTE	900
Db	841	VLDWYPEPSDHFWSLLETFNLRTNNI PGSPGSTELEGGFSRQRKKLSFRRTTDKTE	900
QY	901	QPGVSAIAGPRAACAGPSSRQPGPGMGESPSSGSPSSSEDEGPGRSSPLRLVPSS	960
Db	901	QPGVSAIAGPRAACAGPSSRQPGPGMGESPSSGSPSSSEDEGPGRSSPLRLVPSS	960
QY	961	PRPGEPPGGGAPLMEDECKSSDTNPJLSGAFSGYSNIFSWGDSRGQYQYELPRCAPTP	1020
Db	961	PRPGEPPGGGAPLMEDECKSSDTNPJLSGAFSGYSNIFSWGDSRGQYQYELPRCAPTP	1020
QY	1021	SLNINIPSSPERRRGVYESRLDALORQNLRLERLSADANAIVYLQLRQMTLVPAYSA	1080
Db	1021	SLNINIPSSPERRRGVYESRLDALORQNLRLERLSADANAIVYLQLRQMTLVPAYSA	1080
QY	1081	VTTGPGPSTSPILPVSPPTLLTLDLSI,SOVSQFMACEELPPGAPBELPOEGPTTRLSPG	1140
Db	1081	VTTGPGPSTSPILPVSPPTLLTLDLSI,SOVSQFMACEELPPGAPBELPOEGPTTRLSPG	1140
QY	1141	QLGALTSOPLRHGSDPGS	1159
Db	1141	QLGALTSOPLRHGSDPGS	1159

RESULT 5
AAU99167

AC AAU991677

DT 24-SEP-2002 (first entry)

Human ether-a-go-go related protein, HERG.

KW Human; human ether-a-go-go related gene; HERG; KCRI1, long QT syndrome; IQT; sinoatrial node; polymorphic

potassium channel.
KW
XX

XX
XX

[illegible]

XX

XX

XX
XX
/HVVV-) LINTV VANDERBILT

AA Balser JR. George AL.
PT

WPI; 2002-527650/56.

PT Identifying a potass

PT cerebellar cDNA library (KCR1) polypeptide, and determining activity -

PS Claim 8; Page 158-162; 164pp; English.
xx

CC The invention relates to identifying (MI) a compound that modulates
CC biological activity of a potassium channel (PC) by contacting a

cc cloned from a rat cerebellar cDNA library (KCRI), and determining the

CC compound that modulates the biological activity of a complex comprising a

human ether-a-go-go-related gene (HERG) channel polypeptide and a KCRI polypeptide, identifying (M3) a candidate compound as a modulator of KCRI expression, modulating (M4) PC function in a subject, comprising administering to the subject a substance that provides expression of a KCRI-encoding nucleic acid molecule in a cell or tissue, where modulated PC function is desired, screening (M5) for susceptibility to a drug-induced cardiac arrhythmia in a subject, comprising obtaining a biological sample from the subject and detecting a polymorphism of a KCRI gene in the biological sample from the subject, where the presence of the polymorphism indicates the susceptibility of the subject to a drug-induced cardiac arrhythmia, an oligonucleotide pair, where a first oligonucleotide of the pair hybridises to a first portion of a KCRI gene which includes a polymorphism of the KCRI gene, and the second oligonucleotide of the pair hybridises to a second portion of the KCRI gene that is adjacent to the first portion and a set of antisense oligonucleotide primers, suitable for amplifying a portion of a KCRI gene which includes a polymorphism of the KCRI gene. (M1) is useful for identifying a compound that modulates biological activity of PC, especially HERG, for modulating PC function (i.e. modulating HERG activity) in a mammal, by preparing a composition comprising the compound and administering the composition. The compound is useful for treating or preventing long QT syndrome (LQT) and is useful in drug designing. The present sequence represents human HERG.

Sequence 1159 AA;

Query Match 100.0%; Score 6079; DB 23; Length 1159;
Best Local Similarity 100.0%; Pred. No. 0;

Matches 1159; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

OY 1 MPVRGHAAPQNTFLDTIIIRKFEQSKFTIANANAVENCAYIYCDGCELCGYSRAEVM 60
DB 1 MPVRGHAAPQNTFLDTIIIRKFEQSKFTIANANAVENCAYIYCDGCELCGYSRAEVM 60
OY 61 QRPCTCDLHGPRTDORAAOIAOALGAEEKVYIATYRKDGSCFLCVDVVPVKNEG 120
DB 61 QRPCTCDLHGPRTDORAAOIAOALGAEEKVYIATYRKDGSCFLCVDVVPVKNEG 120
OY 121 AVIMEILNFEYVMEKDMVGSAPADTNRGPTSWLAPRAKTFRLKLPALLATARESSV 180
DB 121 AVIMEILNFEYVMEKDMVGSAPADTNRGPTSWLAPRAKTFRLKLPALLATARESSV 180
OY 181 RSGGAGGAGAGVAVVDLTPAAPSSSLADEVTANDNHVACIGPEERRALVGPSP 240
DB 181 RSGGAGGAGAGVAVVDLTPAAPSSSLADEVTANDNHVACIGPEERRALVGPSP 240
OY 241 PRSAPGOLPSRAHSLNDASGSSCSLARTSRRESCAVRRASSADDIEMRAGVLPSP 300
DB 241 PRSAPGOLPSRAHSLNDASGSSCSLARTSRRESCAVRRASSADDIEMRAGVLPSP 300
OY 301 RHASTGAMHPLRSGILNSTSDSLVRYRTISKIPQITLNFYDLKGPFLASPTSDREIT 360
DB 301 RHASTGAMHPLRSGILNSTSDSLVRYRTISKIPQITLNFYDLKGPFLASPTSDREIT 360
OY 361 PKIKERTHNTEKTYQVLSIGADVLPKYLQAPRIHRTIILHSPFKAVWMLILLVY 420
DB 361 PKIKERTHNTEKTYQVLSIGADVLPKYLQAPRIHRTIILHSPFKAVWMLILLVY 420
OY 421 TAVFTPYSAFLKETEGEPATEGACOPPLAVVDLVDIMFYDILINFRITYVANE 480
DB 421 TAVFTPYSAFLKETEGEPATEGACOPPLAVVDLVDIMFYDILINFRITYVANE 480
OY 481 EVVSHPRGRIAVHYFKGMFLIDMVAIPFDLLIFSGSEELIGLKTARLLRVYARKLD 540
DB 481 EVVSHPRGRIAVHYFKGMFLIDMVAIPFDLLIFSGSEELIGLKTARLLRVYARKLD 540
OY 541 RYSEYGAVALFLMCTFALLIHMILACTIYALGNNEQPMDSRIGLHMLGQIGKPYSS 600
DB 541 RYSEYGAVALFLMCTFALLIHMILACTIYALGNNEQPMDSRIGLHMLGQIGKPYSS 600
OY 601 GLGSPSTKDKVYVLTFTFSSLTSGFNGVSPNTNSEKIFISICVALLIGSLMAYASIFGVS 660
DB 601 GLGSPSTKDKVYVLTFTFSSLTSGFNGVSPNTNSEKIFISICVALLIGSLMAYASIFGVS 660

```

```

OY 661 AIIQRLVSGTARVHTOMLRVEEFIRFHQIPNPLRQRLBEYFOHANSYINGIDMNAVLKGF 720
DB 661 AIIQRLVSGTARVHTOMLRVREFIRFHQIPNPLRQRLBEYFOHANSYINGIDMNAVLKGF 720
OY 721 PECLADICLHINSLIOHCKPFGATVGCRLALAMKFTTHAPPGDTLVHAGDLLVLY 780
DB 721 PECLADICLHINSLIOHCKPFGATVGCRLALAMKFTTHAPPGDTLVHAGDLLVLY 780
OY 781 FISRSIFILNGDVVVALITGNNDIFGEPDLNIYARPKSGNGVRLATYCDLKHIRDDLE 840
DB 781 FISRSIFILNGDVVVALITGNNDIFGEPDLNIYARPKSGNGVRLATYCDLKHIRDDLE 840
OY 841 VLDMPERSDFHWSLLETFNLBDTNTMPCSPGSTLELGGFSRQKRKLSEFRRTDKDTE 900
DB 841 VLDMPERSDFHWSLLETFNLBDTNTMPCSPGSTLELGGFSRQKRKLSEFRRTDKDTE 900
OY 901 QPGEVSALGPBAGAGPSSRGPBGWGESPPSGSPSESSDEGPGRSSPLRLVPSS 960
DB 901 QPGEVSALGPBAGAGPSSRGPBGWGESPPSGSPSESSDEGPGRSSPLRLVPSS 960
OY 961 PRPEEPGPGPIMDECKESSDTCNPLSGATSVNIFSPWGDSDRGROYELPCPCAPT 1020
DB 961 PRPEEPGPGPIMDECKESSDTCNPLSGATSVNIFSPWGDSDRGROYELPCPCAPT 1020
OY 1021 SLNITPLSSPGRPRGDESRDLALQRLNLETRLSDAMATVQLQRLQMTLVPPAYSA 1080
DB 1021 SLNITPLSSPGRPRGDESRDLALQRLNLETRLSDAMATVQLQRLQMTLVPPAYSA 1080
OY 1081 VTTGPGGTSTSPLLPVSPPLTLITDSQVSOQFACCELPAGBELPQEGPTRLSLPG 1140
DB 1081 VTTGPGGTSTSPLLPVSPPLTLITDSQVSOQFACCELPAGBELPQEGPTRLSLPG 1140
OY 1141 QLGALTSOPLHRHGSPPGS 1159
DB 1141 QLGALTSOPLHRHGSPPGS 1159

```

```

RESULT 6
ID ABP71697 standard; Protein: 1159 AA.
XX ABP71697;
AC 12-JUN-2003 (first entry)
DT 12-JUN-2003 (first entry)
XX
DE Human HERG potassium channel polypeptide.
XX
KW HERG; cardiotoxicity; cardiac arrhythmia; ether-a-go-go-related gene;
KW potassium channel; astemizole; human.
OS Homo sapiens.
PN WO2003006988-A1.
XX
PD 23-JAN-2003.
XX
PF 02-JUL-2002; 2002WO-EP07364.
XX
PR 13-JUL-2001; 2001EP-0202689.
XX
PA (JANC ) JANSSEN PHARM NV.
XX
PI Heylen GICM, Janssen CGM, Jurzak M, Van Assouw HPM;
XX
DR WPT: 2003-267961/26.
XX
DR N-PSDB: ABZ76227.
XX
PT Screening test compounds to induce cardiotoxicity/cardiac arrhythmia,
PT by incubating human ether-a-go-go-related gene channel, with a
PT reference and test compound, and measuring effect of the test on
PT reference compound -
XX

```

PS Claim 17; Page 27-29; 43pp; English.
 XX
 CC The invention relates to an assay for screening test compounds for their
 CC capability to induce cardiotoxicity or cardiac arrhythmia in a subject.
 CC The method involves incubating a source containing human ether-a-go-go-
 CC related gene channel (HERG) or a fragment of it with a reference compound
 CC and a test compound, and measuring the effect of the test compound on the
 CC amount of reference compound bound to HERG. A radiolabeled astemizole is
 CC useful as the reference compound in the method for screening test
 CC compounds for their capability to induce cardiotoxicity or cardiac
 CC arrhythmia in a subject. The present sequence represents a human
 CC HERG potassium channel polypeptide.
 XX
 XX Sequence 1159 AA;
 SQ
 Query Match 100.0%; Score 6079; DB 24; Length 1159;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1159; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MPVRGHVAPONTFLDTIIRKEGOSRKEFIANARVENCAYICNDPCELCGYSRAEYV 60
 DB 1 MPVRGHVAPONTFLDTIIRKEGOSRKEFIANARVENCAYICNDPCELCGYSRAEYV 60
 QY 61 QRPCTDFLHGPTORRAAQAIAALLGAEBKVEIAFYRKDGSCFLCIVDVPVKNEG 120
 DB 61 QRPCTDFLHGPTORRAAQAIAALLGAEBKVEIAFYRKDGSCFLCIVDVPVKNEG 120
 QY 121 AVIMFLNFEVMEKMGSPADHTNHRGPTSMALPGRAKPTRLPALALTARESSV 180
 DB 121 AVIMFLNFEVMEKMGSPADHTNHRGPTSMALPGRAKPTRLPALALTARESSV 180
 QY 181 RSGGAGAGAPGAVVDVUTPAAPSSSESLADEVTAMDNHVAAGLGPAAERRALVPGSP 240
 DB 181 RSGGAGAGAPGAVVDVUTPAAPSSSESLADEVTAMDNHVAAGLGPAAERRALVPGSP 240
 QY 241 PRSAPGQLPSRAHSINPDASGSSCSLARTRESGASVRRASSADIDEAMRAGVLP 300
 DB 241 PRSAPGQLPSRAHSINPDASGSSCSLARTRESGASVRRASSADIDEAMRAGVLP 300
 QY 301 RRASTGAMHPLRSGLNSTSDSLVYRPTISKIPQITLNPVDLKGDPFLASPSDEBITA 360
 DB 301 RRASTGAMHPLRSGLNSTSDSLVYRPTISKIPQITLNPVDLKGDPFLASPSDEBITA 360
 QY 361 PKIKERTHNVTEKVTQVLSLADVLPEYKLOAPRIHRMTILHSPKAWMDLILLV 420
 DB 361 PKIKERTHNVTEKVTQVLSLADVLPEYKLOAPRIHRMTILHSPKAWMDLILLV 420
 QY 421 TAVFTPYSAFLIKETEGRPAATECGYACQPLAVVDLIVDIMEVDILINFRTYYNANE 480
 DB 421 TAVFTPYSAFLIKETEGRPAATECGYACQPLAVVDLIVDIMEVDILINFRTYYNANE 480
 QY 481 EYVSHHGRIAVHYFKGMFLIDWVAALPFDLLIFGSGSEELIGLTKTARLLRYVAR 540
 DB 481 EYVSHHGRIAVHYFKGMFLIDWVAALPFDLLIFGSGSEELIGLTKTARLLRYVAR 540
 QY 541 RYSEYGAALVFLMCFALIAHMLACIWAIGMDEPHMDSRGMWLNHNGDQGRKYNSS 600
 DB 541 RYSEYGAALVFLMCFALIAHMLACIWAIGMDEPHMDSRGMWLNHNGDQGRKYNSS 600
 QY 601 GLGGSPIKDKYVYALYFTFSSLTSGVGNVSPNTSEKIFSIICVMLIGSLMAYSI 660
 DB 601 GLGGSPIKDKYVYALYFTFSSLTSGVGNVSPNTSEKIFSIICVMLIGSLMAYSI 660
 QY 661 AIIORLYSGTARYHOMLVRERIRHOIPNPLROLEVEFOHAMSYTGIDMNAVLKGF 720
 DB 661 AIIORLYSGTARYHOMLVRERIRHOIPNPLROLEVEFOHAMSYTGIDMNAVLKGF 720
 QY 721 PECLQADICLHLNRSLLQCKPFGATKGLRLAKFKTTNAPPDPTLVHAGDILLTALY 780
 DB 721 PECLQADICLHLNRSLLQCKPFGATKGLRLAKFKTTNAPPDPTLVHAGDILLTALY 780
 QY 781 FLSRGSIEILRGDVVAAILGKNDIFGEPLNLVAPGKSGDVALTYCDLHKIHRD 840
 DB 781 FLSRGSIEILRGDVVAAILGKNDIFGEPLNLVAPGKSGDVALTYCDLHKIHRD 840

DB 781 FLSRGSIEILRGDVVAAILGKNDIFGEPLNLVAPGKSGDVALTYCDLHKIHRD 840
 QY 841 VLDMPEDSFHFWSSLEITFNLNLTNMIGSPGSTELEGGFSQRRKRKLSFRRTDKOTE 900
 DB 841 VLDMPEDSFHFWSSLEITFNLNLTNMIGSPGSTELEGGFSQRRKRKLSFRRTDKOTE 900
 QY 901 QPGEVSLGPRGAGAPSSSRGPRGPGWGESPPSSGPESSSEDEGGRSSPLRLVPFSS 960
 DB 901 QPGEVSLGPRGAGAPSSSRGPRGPGWGESPPSSGPESSSEDEGGRSSPLRLVPFSS 960
 QY 961 PRPGEPGPEPLMECEKSSDPCNLSGAFSVSNIFSFWGRSRRQVLEPRCAPPT 1020
 DB 961 PRPGEPGPEPLMECEKSSDPCNLSGAFSVSNIFSFWGRSRRQVLEPRCAPPT 1020
 QY 1021 SLNITPLSSPRGRPRDVEDSRDALROLNRLTETRSADMAVYLOLOROMTLVPPAYSA 1080
 DB 1021 SLNITPLSSPRGRPRDVEDSRDALROLNRLTETRSADMAVYLOLOROMTLVPPAYSA 1080
 QY 1081 VTTPGPGPTSTPLPLVSPPLTLTLDLSQVSOFMACEELPPGAPELPQGPTRRLSLPG 1140
 DB 1081 VTTPGPGPTSTPLPLVSPPLTLTLDLSQVSOFMACEELPPGAPELPQGPTRRLSLPG 1140
 QY 1141 QLGALTSQPLHRHSDPGS 1159
 DB 1141 QLGALTSQPLHRHSDPGS 1159
 RESULT 7
 ABP71255
 ID ABP71255 standard; Protein; 1159 AA.
 XX
 AC ABP71255;
 XX
 DT 28-APR-2003 (first entry)
 DE
 DE Human HERG polypeptide.
 XX
 KW Teleost; zebrafish; ZERG; cardiovascular; antiarrhythmic; cytostatic;
 XX neuroprotective; gene therapy; potassium channel; human; HERG.
 OS
 OS Brachydanio rerio.
 PN
 PN WO2003006502-A2.
 PD
 PD 23-JAN-2003.
 PF
 PF 11-JUL-2002; 2002WO-IB04280.
 PR
 PR 13-JUL-2001; 2001US-305396P.
 PA (ARTE-) ARTEMIS PHARM GMBH.
 XX
 XX Langheirrich U;
 PI
 PI
 DR
 DR WPI; 2003-210421/20.
 XX
 PT New teleost (specifically zebrafish) ERG genes, which encode ERG family
 PT potassium channels, useful for studying e.g. cardiac or brain function,
 PT or for developing treatments for e.g. cardiac diseases, cancer or
 PT neurological diseases -
 PS
 PS Examples; Page 50-55; 55pp; English.
 CC
 CC The invention relates to a teleost (zebrafish) ERG (ZERG) polypeptide.
 CC The ZERG nucleic acid molecule is useful for studying cardiac function,
 CC abnormal heart beat phenotype, or long QT syndrome (an abnormality of
 CC cardiac muscle repolarization that predisposes affected individuals to
 CC lethal arrhythmias). The zebrafish comprising ZERG gene is useful as
 CC models for cardiac function or disease. The ZERG genes are particularly
 CC useful for in (non-)cardiac researches, or for developing treatments for
 CC cardiac diseases, tumours or cancers, brain and nervous system disorders
 CC or neurological diseases, or insulin-related diseases. The present
 CC sequence represents a human HERG (ether-a-go-go-related) polypeptide.

```

XX Sequence      1159 AA:
SQ
Query Match      100.0%; Score 6079; DB 24; Length 1159;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1159; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MPVRGHVAPONTFLDITIRKFEQSRKFTIANARVENCAYICNDGFCGLCYSRAEVM 60
DB 1 MPVRGHVAPONTFLDITIRKFEQSRKFTIANARVENCAYICNDGFCGLCYSRAEVM 60
QY 61 ORPCTCFILHGPRTORRAAOIAQALLGAERKVEIAFYRKDSCFCFLCVVYVVKKNDG 120
DB 61 ORPCTCFILHGPRTORRAAOIAQALLGAERKVEIAFYRKDSCFCFLCVVYVVKKNDG 120
QY 121 AVIMFILNFEVMEKDMWGSFAHDTNHRGPTSMILAPRAKTFRLKLPALLALTARESSV 180
DB 121 AVIMFILNFEVMEKDMWGSFAHDTNHRGPTSMILAPRAKTFRLKLPALLALTARESSV 180
QY 121 AVIMFILNFEVMEKDMWGSFAHDTNHRGPTSMILAPRAKTFRLKLPALLALTARESSV 180
DB 121 AVIMFILNFEVMEKDMWGSFAHDTNHRGPTSMILAPRAKTFRLKLPALLALTARESSV 180
QY 181 RSGGAGAGAGAVVVDVLTLPAPSSSESLADEVTAMDHNHAGLGAEBERRALVPGSP 240
DB 181 RSGGAGAGAGAVVVDVLTLPAPSSSESLADEVTAMDHNHAGLGAEBERRALVPGSP 240
QY 241 PRSAPGOLSPRAHSINPDAGSSCSLARTSRSCASVRASSADDIEMARRAGVLP 300
DB 241 PRSAPGOLSPRAHSINPDAGSSCSLARTSRSCASVRASSADDIEMARRAGVLP 300
QY 301 RHAATGAMHPLRSLNSTSDSLVRYRTISKIPQITLNFYDLKGDPFLASPTSDREITIA 360
DB 301 RHAATGAMHPLRSLNSTSDSLVRYRTISKIPQITLNFYDLKGDPFLASPTSDREITIA 360
QY 361 PKIKERHNTVEKYTOVLISGADVLPEYKIQAPRIHMTILHYSFPAVMDLILLVYI 420
DB 361 PKIKERHNTVEKYTOVLISGADVLPEYKIQAPRIHMTILHYSFPAVMDLILLVYI 420
QY 421 TAVFTPSAFLKETEGPRATECGYACOPLAVDLVDIMFTVDLILNRTYVANE 480
DB 421 TAVFTPSAFLKETEGPRATECGYACOPLAVDLVDIMFTVDLILNRTYVANE 480
QY 481 EVVSHPRIAVHYFKGFLIDMVAIPFDLIFGSSSEELIGLKTARLLVAVARKLD 540
DB 481 EVVSHPRIAVHYFKGFLIDMVAIPFDLIFGSSSEELIGLKTARLLVAVARKLD 540
QY 541 RYSEYGAUVEFLMCTALLAHMLACTIWAIGNMEOPHMDRIGMLHNGDQICKPYNS 600
DB 541 RYSEYGAUVEFLMCTALLAHMLACTIWAIGNMEOPHMDRIGMLHNGDQICKPYNS 600
QY 601 GLGSPSTKDKYVTALEYFESSITSVGNGVSPNTNSEKIFISICVMLIGSLMYASIFGVNS 660
DB 601 GLGSPSTKDKYVTALEYFESSITSVGNGVSPNTNSEKIFISICVMLIGSLMYASIFGVNS 660
QY 661 AIIORLSTGATRYTQMLRVREFTIRFHOIPNPLRQRLSEYFOHAMSTNGIDMAVAKGF 720
DB 661 AIIORLSTGATRYTQMLRVREFTIRFHOIPNPLRQRLSEYFOHAMSTNGIDMAVAKGF 720
QY 721 PECLQADICHLNLSLQHCKPRFGATKGCILALAMFKTTNADPGDTLVHAGDLTALY 780
DB 721 PECLQADICHLNLSLQHCKPRFGATKGCILALAMFKTTNADPGDTLVHAGDLTALY 780
QY 781 FISGSEILIRGDNVVAIIGKNDIFGEPMLNYPARKSNGVYALTYCDLHKIRDDILE 840
DB 781 FISGSEILIRGDNVVAIIGKNDIFGEPMLNYPARKSNGVYALTYCDLHKIRDDILE 840
QY 841 VLDNYPEFSDHFWMSLEITFNLRTNMIPGSPSTELGEGFSRORRKLFRRTDQTE 900
DB 841 VLDNYPEFSDHFWMSLEITFNLRTNMIPGSPSTELGEGFSRORRKLFRRTDQTE 900
QY 901 QPGEVSALGPRAGAGPSSRGPRGPGWESBSSPSSPESEDEGPGRSSSPLRLVFPSS 960
DB 901 QPGEVSALGPRAGAGPSSRGPRGPGWESBSSPSSPESEDEGPGRSSSPLRLVFPSS 960
QY 961 PRPGEPPGEPPLMEDCEKSSDTCNPLSGAFSGVSNIFSWGDSRGROYELPRCPAPTP 1020
DB 961 PRPGEPPGEPPLMEDCEKSSDTCNPLSGAFSGVSNIFSWGDSRGROYELPRCPAPTP 1020

```

```

DB 961 PRPGEPPGEPPLMEDCEKSSDTCNPLSGAFSGVSNIFSWGDSRGROYELPRCPAPTP 1020
QY 1021 SLNLIPLSSPPRRRGDVESRLDALQRLNLELTRLSADMATVQLLQROMTLVPPAYSA 1080
DB 1021 SLNLIPLSSPPRRRGDVESRLDALQRLNLELTRLSADMATVQLLQROMTLVPPAYSA 1080
QY 1081 VTPPGPGPTSTSPILLPVSPLETLTLDLSISQVSQFMACEELPPGAPBELPQEGPTRLSLPG 1140
DB 1081 VTPPGPGPTSTSPILLPVSPLETLTLDLSISQVSQFMACEELPPGAPBELPQEGPTRLSLPG 1140
QY 1141 QLGAITSQPLHRHGSDPGS 1159
DB 1141 QLGAITSQPLHRHGSDPGS 1159

RESULT 8
AAY85406
ID AAY85406 standard; Protein: 1159 AA.
AC AAY85406;
XX 19-JUN-2000 (first entry)
DT
DE Long QT syndrome associated HERG protein.
XX
KW HERG; mutation; long QT syndrome; LQT syndrome; gene therapy; human.
XX
OS Homo sapiens.
XX
EN WO200006772-A1.
XX
PD 10-FEB-2000.
XX
PF 20-JUL-1999; 99WO-US16337.
XX
PR 27-JUL-1998; 98US-0122847.
XX
PR 06-JAN-1999; 99US-0226012.
XX
PA (UTAH ) UNIV UTAH RES FOUND.
XX
PI Keating MT, Splawski I;
XX
DR WP1: 2000-195319/17.
XX
DR N-PSDB; AAA07602.
XX
PT New isolated mutant HERG nucleic acids, useful for developing products
PS for the diagnosis, prevention and treatment of long QT syndrome -
PS Disclosure: Fig 8A-B; 163pp; English.
XX
CC The invention relates to a HERG protein having a mutation compared to
CC wild-type HERG, and is useful for developing products for the diagnosis,
CC prevention and treatment of long QT (LQT) syndrome. The products and
CC methods can be used for the diagnosis of subjects with LQT syndrome. They
CC can also be used to screen for drugs for treating or preventing LQT
CC syndrome. The HERG nucleic acids can also be used for gene therapy and
CC HERG peptides can be used for peptide therapy. The present sequence
CC represents the LQT syndrome associated HERG protein.
XX
SQ Sequence 1159 AA:
Query Match      99.9%; Score 6072; DB 21; Length 1159;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1158; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MPVRGHVAPONTFLDITIRKFEQSRKFTIANARVENCAYICNDGFCGLCYSRAEVM 60
DB 1 MPVRGHVAPONTFLDITIRKFEQSRKFTIANARVENCAYICNDGFCGLCYSRAEVM 60
QY 61 ORPCTCFILHGPRTORRAAOIAQALLGAERKVEIAFYRKDSCFCFLCVVYVVKKNDG 120
DB 61 ORPCTCFILHGPRTORRAAOIAQALLGAERKVEIAFYRKDSCFCFLCVVYVVKKNDG 120

```

```

QY 121 AVIMFLNEVNEKDMVSPADHTNHRGPTSMALPGRAKTERLKLPALLATARESSV 180
    |||
Db 121 AVIMFLNEVNEKDMVSPADHTNHRGPTSMALPGRAKTERLKLPALLATARESSV 180
QY 181 RSGGAGAGAGAVVVDVLTTPAAPSESLALDEVMTAMDNHVGGLGPAEERRALVGPSP 240
    |||
Db 181 RSGGAGAGAGAVVVDVLTTPAAPSESLALDEVMTAMDNHVGGLGPAEERRALVGPSP 240
QY 241 PRSAPOLSPRRASHINLPDASGSSCSLARPRRESKASVRRASSADIEMRAGVLP 300
    |||
Db 241 PRSAPOLSPRRASHINLPDASGSSCSLARPRRESKASVRRASSADIEMRAGVLP 300
QY 301 RAHSTGAMHPLRSGLNSTSDSLVYRTSKIPQITLNVVDKGFPLSPSDREIIA 360
    |||
Db 301 RAHSTGAMHPLRSGLNSTSDSLVYRTSKIPQITLNVVDKGFPLSPSDREIIA 360
QY 361 PKIKERTHNVEKVTQVLSGADVLDEYKLAQAPRIHRMTLHYSPFKAVMDMLILLV 420
    |||
Db 361 PKIKERTHNVEKVTQVLSGADVLDEYKLAQAPRIHRMTLHYSPFKAVMDMLILLV 420
QY 421 TAVFTPYSAFLAKETEEGPATCEGACOPPLAVVDLIVIMFIVDLINERTTYVANE 480
    |||
Db 421 TAVFTPYSAFLAKETEEGPATCEGACOPPLAVVDLIVIMFIVDLINERTTYVANE 480
QY 481 EVVSHPGRIAVHYFKGMFLIDMYAALPFDDLIFGSGSEELIGLKTARLLRLVAVARK 540
    |||
Db 481 EVVSHPGRIAVHYFKGMFLIDMYAALPFDDLIFGSGSEELIGLKTARLLRLVAVARK 540
QY 541 RYSEYGAANVFLMLCTFALIAHMLACTIWAIGNMEOPHMDSRIGWLNHLDQIGKPN 600
    |||
Db 541 RYSEYGAANVFLMLCTFALIAHMLACTIWAIGNMEOPHMDSRIGWLNHLDQIGKPN 600
QY 601 GLGGPSTKDKYVYALFTFSSLSYVGNGNSPMTNSKITSICVMLIGSLMYISFGN 660
    |||
Db 601 GLGGPSTKDKYVYALFTFSSLSYVGNGNSPMTNSKITSICVMLIGSLMYISFGN 660
QY 661 AIIOQRTYSGTARYTQMLRVREFIRFHQINPPLRQLEEFYOHAMSTYNGIDMNAV 720
    |||
Db 661 AIIOQRTYSGTARYTQMLRVREFIRFHQINPPLRQLEEFYOHAMSTYNGIDMNAV 720
QY 721 PECLQADICLHLNRSLLQHCPPRGATKGLRALAMKFKTTTHAPPGDVLVHAGD 780
    |||
Db 721 PECLQADICLHLNRSLLQHCPPRGATKGLRALAMKFKTTTHAPPGDVLVHAGD 780
QY 781 F1RSGSTEILRGDVVAAILGKNDIFGEPMLYARPGKSNQDVNALTCDLHKIHRD 840
    |||
Db 781 F1RSGSTEILRGDVVAAILGKNDIFGEPMLYARPGKSNQDVNALTCDLHKIHRD 840
QY 841 VLDMPPEFSDHFMSSLEITFNLDNTNMIPGSPSTEEGGFSRQRRKLSFRRTD 900
    |||
Db 841 VLDMPPEFSDHFMSSLEITFNLDNTNMIPGSPSTEEGGFSRQRRKLSFRRTD 900
QY 901 QPEVSALGPGRAGAGPSNRGPGGPGWGESPPSSGPESEDEGPGRSSSPLRV 960
    |||
Db 901 QPEVSALGPGRAGAGPSNRGPGGPGWGESPPSSGPESEDEGPGRSSSPLRV 960
QY 961 PRPPGEPGGEPLMEDECKSSDTCNPLSGAFSGVSNFISWGSRGQYDELPRCP 1020
    |||
Db 961 PRPPGEPGGEPLMEDECKSSDTCNPLSGAFSGVSNFISWGSRGQYDELPRCP 1020
QY 1021 SLNINIPSSGRRPRGVDVESRLDALQRLNLETRLSADAMATVQLLQORONT 1080
    |||
Db 1021 SLNINIPSSGRRPRGVDVESRLDALQRLNLETRLSADAMATVQLLQORONT 1080
QY 1081 VTPPGGPTSTSPLLPVSPPTLTLDLSQVSOFMACEELPQGAPELPQEGPFRRL 1140
    |||
Db 1081 VTPPGGPTSTSPLLPVSPPTLTLDLSQVSOFMACEELPQGAPELPQEGPFRRL 1140
QY 1141 QLGALTSQPLHRHSGDPGS 1159
    |||
Db 1141 QLGALTSQPLHRHSGDPGS 1159

```

```

RESULT 9
AA83031
ID AAV83031 standard; Protein: 1163 AA.
XX
AC AAV83031;
XX
DT 04-JUL-2000 (first entry)
XX
DE Rat Erg1 potassium channel membrane protein.
XX
KW Elk1; elk2; eag1; eag2; erg1; potassium channel; membrane protein;
KW drug screening; hypertension; renal failure; diabetes insipidus;
KW diabetic nephropathy; hypothyroidism; goiter; hypoparathyroidism;
KW pancreatic insufficiency; diabetes mellitus; cystic fibrosis;
KW salivary; salivary insufficiency; diabetes mellitus; current flow;
KW ion flux; transcription; signal transduction; assay; detection;
KW rat.
XX
OS Rattus rattus.
XX
PN W0200012546-A1.
XX
PD 09-MAR-2000.
XX
PE 31-AUG-1999; 99WO-US19902.
XX
PR 31-AUG-1998; 98US-0098413.
XX
PA (UYN ) UNIV NEW YORK STATE RES FOUND.
XX
PI McKinnon D, Dixon JE;
XX
DR WPI: 2000-256585/22.
XX
PT Novel mammalian potassium channel genes and polypeptides encoded by
PT them for screening drugs useful for treating diseases such as
PT hypertension, acute renal failure, diabetes insipidus and
PT hypothyroidism -
XX
PS Disclosure; Page 90-95; 102pp; English.
XX
CC Potassium channel genes e.g. elk1, elk2 or eag2 are useful for
CC identifying modulators which are useful for treating hypertension,
CC acute and chronic renal failure, diabetes insipidus, diabetic
CC nephropathy, hyperthyroidism, goiter, hyperparathyroidism,
CC pancreatic insufficiency, diabetes mellitus, cystic fibrosis,
CC salivary, salivary insufficiency. The availability of the gene
CC sequences provides a tool for research into the physiological
CC characteristics of the various genes and proteins for potassium
CC channels including the development of medicines effective for
CC treating disease conditions associated with mutations or defects in
CC potassium channels and the screening of drugs to ensure that
CC potassium channels are not blocked or physiologically affected by
CC those drugs. The channel proteins encoded by these genes are also
CC useful to measure changes in reporter molecules in assay and detection
CC systems to measure changes in potassium concentration, membrane
CC potential, current flow, ion flux, transcription, signal
CC transduction, receptor-ligand interaction and second messenger
CC concentrations. See also GENESEQ records AA293334-293336.
XX
XX
XX Sequence 1163 AA:
XX
Query Match 95.1%; Score 5782; DB 21; Length 1163;
Best Local Similarity 95.3%; Pred. No. 0;
Matches 1111; Conservative 12; Mismatches 36; Indels 4; Gaps 2;
QY 1 MPVARGHVAPONFTLDTIIRFEGOSRKFIIANARVENCAVIYCNDFCELCGSR 60
    |||
Db 1 MPVARGHVAPONFTLDTIIRFEGOSRKFIIANARVENCAVIYCNDFCELCGSR 60
QY 61 QRPCTCDFLHGPTQRRRAAAQIAQALLGAERKVEIAFYRRDGSCTCLVDVVPVK 120
    |||
Db 61 QRPCTCDFLHGPTQRRRAAAQIAQALLGAERKVEIAFYRRDGSCTCLVDVVPVK 120

```

```

OY 121 AVIMFLNEFVMEKDMGSPAHDTNHRGPTSWLAPGRAKTFRLKLPALLALTARESSV 180
DB 121 AVIMFLNEFVMEKDMGSPAHDTNHRGPTSWLMSGAKTFRLKLPALLALTARESPM 180
OY 181 RSGGAGAGAPAAVVVDVLTLPAAAPSSSLADEVTAMDNHVAIGPAPERRALVGP--G 238
DB 181 RSGTSGPAGAAVVVDVLTLPAAAPSSSLADEVTAMDNHVAIGPAPERRALVGPASA 240
OY 239 SPBRAPGOLPSPRAHSLNPDAGSSCSLARTRESGASVARRASADDIEAMRAGVLP 298
DB 241 SPVASIPGHPSPRAQNSLPDAGSSCSLARTRESGASVARRASADDIEAMRAGALPL 300
OY 299 PPRHASTGAMHLRSGGLNSTSDSLVYKRTISKIPQITLNVLDKGPFLASPTSDREI 358
DB 301 PPRHASTGAMHLRSGGLNSTSDSLVYKRTISKIPQITLNVLDKGPFLASPTSDREI 360
OY 359 IAPKIKERHNTVEYTOVLSGADVLPDYKLOARIRHWTILHSPKFAVDMWILLLV 418
DB 361 IAPKIKERTHNTVEYTOVLSGADVLPDYKLOARIRHWTILHSPKFAVDMWILLLV 420
OY 419 IYTAFTPEYSAFLLKETEEGPATEGACOPLAIVVDLIVDIMEIVLILNFTTYNA 478
DB 421 IYTAFTPEYSAFLLKETEDGSOADCGYACQPLAVVDLIVDIMEIVLILNFTTYNA 480
OY 479 NEEVYSHRGRIAVHFKGMFLIDMAAIPFDLLIFGSGSEELIGLKTARLLRLYVARK 538
DB 481 NEEVYSHRGRIAVHFKGMFLIDMAAIPFDLLIFGSGSEELIGLKTARLLRLYVARK 540
OY 539 IDRYSEYGAAYVFLMCTFEALIAHMLACIWAIGMBOPHMDSRIGMHNIDOGKPYN 598
DB 541 IDRYSEYGAAYVFLMCTFEALIAHMLACIWAIGMBOPHMDSRIGMHNIDOGKPYN 600
OY 599 SSGLGSPITKDYVALYFETSSLSVGRGVNSPTNSEKIFPSICVMLIGSLMTASIRGN 658
DB 601 SSGLGSPITKDYVALYFETSSLSVGRGVNSPTNSEKIFPSICVMLIGSLMTASIRGN 660
OY 659 VSAIIQRIYSGTARHTOMLRRETIREFHOIIPNPRORLEEFQHAMSTNGIDMNAVILK 718
DB 661 VSAIIQRIYSGTARHTOMLRRETIREFHOIIPNPRORLEEFQHAMSTNGIDMNAVILK 720
OY 719 GPPELQADICLHLNRSLLOHCKPRGATKGLRALAMKFKTTNHPGDTLVHAGDLTLA 778
DB 721 GPPELQADICLHLNRSLLOHCKPRGATKGLRALAMKFKTTNHPGDTLVHAGDLTLA 780
OY 779 LYFISRGSIELRGDVVAALIGKNDIFGEPLNLYARPGKNSGDVYALTYCDLHKIHRDL 838
DB 781 LYFISRGSIELRGDVVAALIGKNDIFGEPLNLYARPGKNSGDVYALTYCDLHKIHRDL 840
OY 839 LEVIDMYPEFSDHFWSSLEITFNLRDTNMTIPGSPGSTELEGFSQORRKLISFRRTOKD 898
DB 841 LEVIDMYPEFSDHFWSSLEITFNLRDTNMTIPGSPGSTELEGFSQORRKLISFRRTOKD 900
OY 899 TEQGEVYAL--GPRAGAGPSSRCRPGGPGWGESPPSSPESSEDDGPGSSSPLV 956
DB 901 TEQGEVYALOGCARVPGPSCRCQPGGPGWGESSSPSPESSEDDGPGSSSPLV 960
OY 957 PFSSRPPEGPGEPLMEDCEKSSDTCNPLSGAFSGVNTIFSEWDSRGRQYOLPRCP 1016
DB 961 PFSSRPPEGPGEPLTEDEKSSDTCNPLSGAFSGVNTIFSEWDSRGRQYOLPRCP 1020
OY 1017 APTBSLNLPISSGERRRGVYESKDALQROLNLETRLASADATVYQLQROMTLVLP 1076
DB 1021 AAPASLNLPISSGERRRGVYESKDALQRODNLETRLASADATVYQLQROMTLVLP 1080
OY 1077 AYSATVTPGPGPTSNPLPVSPPLTLTLDLSIYVSOQMACEBELPGABELPODEPTRL 1136
DB 1081 AYSATVTPGPGPTSTSPPLPVGPVPTLTLDLSIYVSOQVAFEBELPAGABELPODEPTRL 1140
OY 1137 SLPGOLGALTSQPLHRHGSDDPGS 1159
DB 1141 SLPGOLGALTSQPLHRHGSDDPGS 1163

```

```

RESULT 10
ABP71254
ID ABP71254 standard; Protein; 1177 AA.
XX
AC ABP71254;
XX
DE 28-APR-2003 (first entry)
XX
DE Chimeric ZERG polypeptide.
XX
KM Teleost; zebrafish; ZERG; cardiovascular; antiarrhythmic; cytostatic;
KM neuroprotective; gene therapy; potassium channel; chimeric..
OS Brachydanio rerio.
XX Homo sapiens.
XX
PN WO2003006502-A2.
XX
PD 23-JAN-2003.
XX
PF 11-JUL-2002; 2002WO-IB04280.
XX
PR 13-JUL-2001; 2001US-305396P.
XX
PA (ARTE-) ARTEMIS PHARM GMBH.
XX
PI Langheirnich U;
XX
DR WPI; 2003-210421/20.
XX
PT New teleost (specifically zebrafish) ERG genes, which encode ERG family
PT potassium channels, useful for studying e.g. cardiac or brain function,
PT or for developing treatments for e.g. cardiac diseases, cancer or
PT neurological diseases -
XX
PS Claim 24; Page 46-50; 55pp; English:
XX
CC The invention relates to a teleost (zebrafish) ERG (ZERG) polypeptide.
CC The ZERG nucleic acid molecule is useful for studying cardiac function,
CC abnormal heart beat phenotype, or long QT syndrome (an abnormality of
CC cardiac muscle repolarization that predisposes affected individuals to
CC lethal arrhythmias). The zebrafish comprising ZERG gene is useful as
CC models for cardiac function or disease. The ZERG genes are particularly
CC useful for in (non-)cardiac researches, or for developing treatments for
CC cardiac diseases, tumours or cancers, brain and nervous system disorders
CC or neurological diseases, or insulin-related diseases. The present
CC sequence represents a chimeric zebrafish ZERG polypeptide.
XX
SQ Sequence 1177 AA;

```

Query Match 57.5%; Score 3493.5; DB 24; Length 1177;
Best Local Similarity 61.1%; Pred. No. 4.1e-275;
Matches 757; Conservative 107; Mismatches 227; Indels 147; Gaps 32;

```

OY 1 MPVRRGHVAPONTFLDTIIRKFEQSGKFTIIMARVENCAYIYNDGFCBLGYSRAEVM 60
DB 1 MPVRRGHVAPONTFLDTIIRKFEQSGKFTIIMARVENCAYIYNDGFCBLGYSRAEVM 60
OY 61 QRCPCDFLHPRIQRRRAADLAQALLGAERKVEIATYRRDSCFCLYVVPKKNEDG 120
DB 61 QRCPCDFLHPRIQRRRAADLAQALLGAERKVEIATYRRDSCFCLYVVPKKNEDG 120
OY 61 QRCPCDFLHPRIQRRRAADLAQALLGAERKVEIATYRRDSCFCLYVVPKKNEDG 120
DB 61 QRCPCDFLHPRIQRRRAADLAQALLGAERKVEIATYRRDSCFCLYVVPKKNEDG 120
OY 121 AVIMFLNEFVMEKDMV--GSPAHDTNHRGPTSWLAPGRAKTFRLKLPALLALTARES 178
DB 121 AVIMFLNEFVMEKDMV--GSPAHDTNHRGPTSWLAPGRAKTFRLKLPALLALTARES 178
OY 179 SVRSAGAGAGAPAAVVVDVLTLPAAAPSSSLADEVTAMDNHVAIGLP---AEERRAL 234
DB 179 SVRSAGAGAGAPAAVVVDVLTLPAAAPSSSLADEVTAMDNHVAIGLP---AEERRAL 234
OY 173 SLPKDHFEG-----VVVDY----LQPSHEVALKLDI-----MSRDSCLKSENGAL 214
DB 235 V--GPGSPRRAPGOLPSPRAHSLNPDAGSSCSLARTRESGASVARRASADDIEAMR 292

```



```

Db      215 IQQTPSSCELSPPSPKSDRLPESGP-----LKHSHSRSMHSLARRASSLHDIQNR 267
QY      293 AGVLPPEPPRHASTGAMHPLRSGLLNSTSDILVRYRTISKIQIQTLENFDLKGDPLASP 352
Db      268 -----DOMSDLKPSNLNSTSSDLKMRHTICRIPOVITSF---GSDRLRPP 310
QY      353 T--SDREILAP-KIKERTHNATEK---VTVLSIGADVLEPYKLOAPRIHRTWILHSPRK 407
Db      311 SPTLEIILAPSKIKDRSQNVSEKVTQVVLISGADVLEPYKLOAPRIHRTWILHSPRK 370
QY      408 AWMDDLILLVYTAFTFYSAFLKEPEEGPATEGACQPLAVDLIVDIMEYDI 467
Db      371 AWMDDLILLVYTAFTFYSAFLKEPEEGPATEGACQPLAVDLIVDIMEYDI 430
QY      468 LINFRTTYNANEVYSHPGRIAVHFKGMFLIDWAAIPDLLIFGSGSEELIGLKTA 527
Db      431 LINFRTTYNANEVYSHPGRIAVHFKGMFLIDWAAIPDLLIFGSGSEELIGLKTA 490
QY      528 RLRLVYRVARKLDYRSEYGAVALFLIMCTFALIAHMACIWAIGMEQPHMDSRIGWH 587
Db      491 RLRLVYRVARKLDYRSEYGAVALFLIMCTFALIAHMACIWAIGMEQPHMDSRIGWH 550
QY      588 NIGDOI GKPNSSGJGPGSIKDKYVYALYFTFSSLTSGPQVNSPNTSEKIFSTICVLI 647
Db      551 NIGDOI GKPNSSGJGPGSIKDKYVYALYFTFSSLTSGPQVNSPNTSEKIFSTICVLI 610
QY      648 GSIWASTIFGNYSATIIORLYSGTARYHQMILVREFIRRHQIIPNLROLEEFQFANXY 707
Db      611 GSIWASTIFGNYSATIIORLYSGTARYHQMILVREFIRRHQIIPNLROLEEFQFANXY 670
QY      708 TNGIDMNAVALKGFPEELQADICLHNLRSLLQCKPKPAGTKCLRALAKFKTTAHPGD 767
Db      671 TNGIDMNAVALKGFPEELQADICLHNLRSLLQCKPKPAGTKCLRALAKFKTTAHPGD 730
QY      768 TIYVHAGDILTAIFYTSRSGIETLRGDVYVAIIGKNDIFGEPLNLVAPKSGNDVATLY 827
Db      731 TIYVHAGDILTAIFYTSRSGIETLRGDVYVAIIGKNDIFGEPLNLVAPKSGNDVATLY 790
QY      828 CDLHKIHRDLDLEVIDMVEPFDHFWSSLEITFNLDTMIRGSPGSETELEGGFSQQRK 887
Db      791 CDLHKIHRDLDLEVIDMVEPFDHFWSSLEITFNLDTMIRGSPGSETELEGGFSQQRK 850
QY      888 KLSFRF-----RTDKDTEQGEVSALCPGRAGAPSSRGPGGPGWGESPPSG-PPSPE 939
Db      851 RHPLRNRNRPDGMDRGMOTYVQPCSPGVNHR-GAIPLSQ-----WDLCSGSPASIS 903
QY      940 SSEDECP---GRS---SSPLRLVPRSSPPRPGEGPEPLMEDCEK-SSDTCNPLSGAFS 992
Db      904 SSEDECP---GRS---SSPLRLVPRSSPPRPGEGPEPLMEDCEK-SSDTCNPLSGAFS 955
QY      993 -----GVSNISFWDGSRGROYELPRCAPPTPLSLNIPLSPPGRPRPGDVESRTDALQ 1046
Db      956 YTAAPLAINISGYSTYLSDRASEYSEQKSSSAVOACIYHHSPCVGDRP-NOLOARLELQ 1014
QY      1047 ROLNRLLETSLADMAVTIQLLOROMTLVPAYSAVT-----TPGPGPTS-----T 1091
Db      1015 SOLNRLLETSLADMAVTIQLLOROMTLVPAYSAVT-----TPGPGPTS-----T 1074
QY      1092 SPULRP-----SPULPITLDSLSQVSO-----FMACEELPPCAPP-----LPQEGP-- 1132
Db      1075 TPTSLQITTDASPGKSPGDVDSLSEKSPDSSLSSGIIHLVASTDTMSMSPETELSVSPSPUL 1134
QY      1133 -----TRRLSLPGOL---GALTSOP-LHRHGSOP 1157
Db      1135 QPGLLCSLRFPSPPLDLSLEGFTLEGSFEIQRHVSDP 1172

```

RESULT 11
 ID AA014189 standard: protein; 1196 AA.
 AC AC
 XX AA014189;

```

DT      03-MAY-2002 (first entry)
XX
DE      Human transporter and ion channel TRICH-6.
XX
KW      Human; transporter and ion channel. TRICH; transport disorder;
KW      neurological disorder; muscle disorder; immunological disorder;
KW      cell proliferative disorder; neuroprotective; noctropic;
KW      cerebroprotective; immunosuppressive; cytostatic; respiratory; muscular;
KW      gene therapy.
XX
OS      Homo sapiens.
XX
PN      W0200204520-A2.
XX
PD      17-JAN-2002.
XX
PE      05-JUL-2001; 2001WO-US21448.
XX
PR      07-JUL-2000; 2000US-216547P.
PR      14-JUL-2000; 2000US-218232P.
PR      21-JUL-2000; 2000US-220112P.
PR      28-JUL-2000; 2000US-221839P.
XX
PA      (INCY-) INCYTE GENOMICS INC.
PI      Raumann BE, Thornton M, Ding L, Yue H, Tang YT, Harland L;
PI      Burford N, Greene BD, Sanjaivala MS, Baughn MR, Yao MG, Yang J;
PI      Patterson C, Gandhi AR, Hatalla AJA, Tribouley CM, Walla NK;
PI      Au-Young J, Walsh RT, Ramkumar J, Lu Y, Lu DAM, Azimzai Y, Lal P;
PI      Elliott VS, Nguyen DB, Xu Y, Sellhammer JJ, Borowsky ML, Khan FA;
PI      Kearney L, Thangavelu K, Das D, Pollicky JL.
XX
DR      WPI: 2002-205969/26.
DR      N-PSDB: AAL44672.
XX
XX      New human transporters and ion channel polypeptides for diagnosing,
XX      treating or preventing transport, neurological, muscle, immunological
XX      and cell proliferative disorders -
XX
PS      Claim 1; Page 151-153; 230pp; English.
XX
CC      The present invention provides the protein and coding sequences of a
CC      number of human transporter and ion channel proteins, designated
CC      TRICH-1-TRICH-32. The sequences can be used in the treatment of
CC      transport, neurological, muscle, immunological and cell proliferative
CC      disorders. The present sequence is a protein of the invention.
XX
SQ      Sequence 1196 AA:

```

Query Match 54.2%; Score 3293; DB 23; Length 1196;
 Best Local Similarity 55.2%; Pred. No. 9, 3e-259;
 Matches 705; Conservative 126; Mismatches 241; Indels 206; Gaps 25;

```

QY      1 MPVRGHVAPQNTFLDTIIRKEGSGRKFLIANARVENCAYIYCNDFCELCGYSRAEYV 60
Db      1 MPVRGHVAPQNTFLDTIIRKEGSGRKFLIANARVENCAYIYCNDFCELCGYSRAEYV 60
QY      61 QRPCTDFLHGPRTPQRAAQAIALGAEERVELAFTRKDSGCLCLVDVVPVKNEDG 120
Db      61 QRPCTDFLHGPRTPQRAAQAIALGAEERVELAFTRKDSGCLCLVDVVPVKNEDG 120
QY      121 AVIMFLINFEVWEKMGVSPAHDTNHRGPTSMLAGRAKTPRLKLPALLATARESSV 180
Db      121 VAMFLINFEVYTDNENATP-----ERNVPLIPITVNRKRFPGRLRVLTYYKQSL 175
QY      181 RSGGAGAGAPGAVVVDVLTTPAAPSSSLADEVTAMDNHVAAGLCPAERRALVPG-- 238
Db      176 PQED-----PDVVVID-----SSKHSDDSVAMKHFKSPKESCPSEADDTKALIQPSK 225
QY      239 SPPRSAPGOL-----PSPRAHSLNPDASGSGSLARTRRSKCSAVRASSADIDEMKRG 294
Db      226 SPLVNISGPLDHSPPKRDWRDLYPDMLQSSSOLSHSRSRSLCSIRRASVHDIEGF--G 283

```

QY	295	VLPP- - - - -	PPHRA	-----	-TGAMHPLR	SSGLNST	SOSD	LYR	FTISK	IPDIT	INFL	VDLK	344
						:	:	:	:	:	:	:	:
					:	:	:	:	:	:	:	:	:
Db	284	VHRKNIFR	DRA	S	E	D	N	G	R	N	K	G	PPR
					:	:	:	:	:	:	:	:	:
					:	:	:	:	:	:	:	:	:
QY	345	GDPFLAS-	-P	T	S	D	R	E	L	I	A	P	K
		:	:	:	:	:	:	:	:	:	:	:	:
		:	:	:	:	:	:	:	:	:	:	:	:
Db	344	TEKNSS	SP	S	P	S	S	D	K	T	I	A	P
		:	:	:	:	:	:	:	:	:	:	:	:
		:	:	:	:	:	:	:	:	:	:	:	:
QY	404	SPFAVMD	W	L	L	L	L	L	L	L	L	L	L
		:	:	:	:	:	:	:	:	:	:	:	:
		:	:	:	:	:	:	:	:	:	:	:	:
Db	404	SPFAVMD	W	L	L	L	L	L	L	L	L	L	L
		:	:	:	:	:	:	:	:	:	:	:	:
		:	:	:	:	:	:	:	:	:	:	:	:
QY	464	I	V	D	L	I	N	E	R	T	T	Y	V
		:	:	:	:	:	:	:	:	:	:	:	:
		:	:	:	:	:	:	:	:	:	:	:	:
Db	464	I	V	D	L	I	N	E	R	T	T	Y	V
		:	:	:	:	:	:	:	:	:	:	:	:
		:	:	:	:	:	:	:	:	:	:	:	:
QY	521	I	G	L	L	T	A	R	L	L	R	V	A
		:	:	:	:	:	:	:	:	:	:	:	:
		:	:	:	:	:	:	:	:	:	:	:	:
Db	523	I	G	L	L	T	A	R	L	L	R	V	A
		:	:	:	:	:	:	:	:	:	:	:	:
		:	:	:	:	:	:	:	:	:	:	:	:
QY	581	S	R	I	C	M	L	N	L	C	O	I	G
		:	:	:	:	:	:	:	:	:	:	:	:
		:	:	:	:	:	:	:	:	:	:	:	:
Db	583	D	K	I	S	M	L	S	I	G	O	I	G
		:	:	:	:	:	:	:	:	:	:	:	:
		:	:	:	:	:	:	:	:	:	:	:	:
QY	640	F	S	I	C	M	L	I	G	S	I	M	A
		:	:	:	:	:	:	:	:	:	:	:	:
		:	:	:	:	:	:	:	:	:	:	:	:
Db	643	F	S	I	C	M	L	I	G	S	I	M	A
		:	:	:	:	:	:	:	:	:	:	:	:
		:	:	:	:	:	:	:	:	:	:	:	:
QY	700	Y	F	O	H	A	M	T	Y	N	G	I	D
		:	:	:	:	:	:	:	:	:	:	:	:
		:	:	:	:	:	:	:	:	:	:	:	:
Db	703	Y	F	O	H	A	M	T	Y	N	G	I	D
		:	:	:	:	:	:	:	:	:	:	:	:
		:	:	:	:	:	:	:	:	:	:	:	:
QY	760	T	T	H	A	P	G	E	T	I	A	L	F
		:	:	:	:	:	:	:	:	:	:	:	:
		:	:	:	:	:	:	:	:	:	:	:	:
Db	763	T	T	H	A	P	G	E	T	I	A	L	F
		:	:	:	:	:	:	:	:	:	:	:	:
		:	:	:	:	:	:	:	:	:	:	:	:
QY	820	G	D	V	A	L	T	A	C	D	L	I	H
		:	:	:	:	:	:	:	:	:	:	:	:
		:	:	:	:	:	:	:	:	:	:	:	:
Db	823	A	D	V	A	L	T	A	C	D	L	I	H
		:	:	:	:	:	:	:	:	:	:	:	:
		:	:	:	:	:	:	:	:	:	:	:	:
QY	877	L	E	G	F	S							

RESULT 12
ABP71253
ID ABP71253 standard; Protein; 1186 AA.
XX

AC	ABP1253;
XX	
DT	28-APR-2003 (first entry)
XX	
DE	Zebrafish ZERG polypeptide.
XX	
KM	Telostei; zebrafish; ZERG; cardiovascular; antiarrhythmic; cytostatic;
XX	neuroprotective; gene therapy; potassium channel.
OS	
XX	Brachydanio rerio.
PN	WO2003006502-A2.
XX	
PD	23-JAN-2003.
XX	
PF	11-JUL-2002; 2002WO-IB04280.
XX	
PR	13-JUL-2001; 2001US-305396P.
XX	
PA	(ARTE-) ARTEMIS PHARM GMBH.
XX	
PI	Langheinrich U;
XX	
DR	MPI: 2003-210421/20.
N-PSDB:	ABZ58872.
PT	New teleost (specifically zebrafish) ERG genes, which encode ERG family
PT	potassium channels, useful for studying e.g. cardiac or brain function,
PT	or for developing treatments for e.g. cardiac diseases, cancer or
PT	neurological diseases -
XX	
PS	Examples: Page 40-45; 55pp; English.
XX	
CC	The invention relates to a teleost (zebrafish) ERG (ZERG) polypeptide.
CC	The ZERG nucleic acid molecule is useful for studying cardiac function,
CC	abnormal heart beat phenotype, or long QT syndrome (an abnormality of
CC	cardiac muscle repolarization that predisposes affected individuals to
CC	lethal arrhythmias). The zebrafish comprising ZERG gene is useful as
CC	models for cardiac function or disease. The ZERG genes are particularly
CC	useful for in (non-)cardiac researches, or for developing treatments for
CC	cardiac diseases, tumours or cancers, brain and nervous system disorders
CC	or neurological diseases, or insulin-related diseased. The present
CC	sequence represents a zebrafish ZERG polypeptide, an ERG family potassium
CC	channel polypeptide.
XX	
SQ	Sequence 1186 AA;
Query Match	53.3%; Score 3241; DB: 24; Length 1186;
Best Local Similarity	57.5%; Pred. No. 1.6e-254;
Matches 717; Conservative 121; Mismatches 252;	Indels 158; Gaps 36
OY	1 MPVRGRHVAPONTFLDITIRKFEQSGRRFTIANAVRNCAVITYCNGFCELCYSRAEVM 60
Db	1 MPVRGRHVALLQNTYTLDITIRKFDGCKRKLIANAOKKCGLIYCNGFCOMFESFAEIM 60
OY	61 QRPCTCDLHGRTQRRAAQAIAQLGAEEKRVELAFYAKDGSCFLCVDVVPVKNEGG 120
Db	61 QGSCTCFVLVGRTKMSALGOLAHLGSEERKEVEILYYSKGSTCARCLDIVIPVKNNEG 120
OY	121 AVIMTLTFEEVYMEDMV--GSPAHDTHNRCPETSLADGRAKTFLKLPLALLATARES 178
Db	121 VVIMTLNFOELDLDSMKKKGLIKORMAN-----SWLRAGORRMRLRMPSTL-RVKROP 172
OY	179 SVRSRGAGAGAGAGVVVDVDTLPAPSSSELALDDVTAMDHNHVAIGLG----AEERRAL 234
Db	173 SLPKDHFEg-----VVVDY-----IQPSHEEVALKDLo-----WSPDSCUKSTOAL 214
OY	235 V--GGSPPRSAPOGLPSRAHSLNPDASGSSCSLARTSRBSCASVSRASSADDIEARK 292
Db	215 IQQTSSCSLSPPRPSPDRLEPSCG-----LKHSHSRRESMHSLRRASSLHDIDGAR 267
OY	293 AGVLPDPRHASTGMHPLRSLNLNSTSDSLVRYRTISKIPQITLTINFYDLKGDFLASP 352

[illegible]

RESULT	13
ABG31253	
ID	ABG31253 standard; Protein: 994 AA.
XX	
AC	ABG31253;
XX	
DT	21-OCT-2002 (first entry)
XX	
DE	Human erg2 (h-erg2) protein splice variant
XX	

XX	Human; erg2; hypotensive; hypertensive; cytosolic; antiferility; nephrotropic; potassium channel inhibitor; hypotension; hypertension; renal failure; benign prostatic hyperplasia; prostate cancer; interility; splice variant; potassium channel.
OS	Homo sapiens.
XX	WO200242417-A2.
PN	30-MAY-2002.
XX	16-NOV-2001; 2001WO-US43490.
XX	20-NOV-2000; 2000US-249981P.
XX	(MERI) MERCK & CO INC.
PA	Folander KL, McKenna EJ, Swanson RJ, Liu Y;
PI	WPI: 2002-583376/62.
DR	N-PSDB; ABK86233.
XX	
PT	New isolated human-erg2 potassium channel subunit, useful for treatment
PT	of hypertension, hypotension, renal failure, benign prostate
PT	hyperplasia, prostate cancer and interility
XX	
PS	Claim 8; Fig 3B; 53pp; English.
XX	
CC	This invention relates to an isolated human erg2 potassium channel
CC	subunit protein. The erg2 protein of the invention is useful for
CC	identifying activators or inhibitors of potassium channels containing
CC	the protein. The erg2 protein is also useful in counter screens for
CC	assays designed to identify activators and inhibitors of other drug
CC	targets. The protein is useful for treating hypotension, hypertension,
CC	renal failure, benign prostatic hyperplasia, prostate cancer, and
CC	infertility. The activators and inhibitors of potassium channels
CC	containing h-erg2 protein, identified using this protein are useful for
CC	treating or preventing conditions as described above, where the
CC	activity of potassium channels containing h-erg2 protein is abnormal.
CC	The nucleic acid encoding the human erg2 protein is useful in
CC	various diagnostic methods, and a DNA or RNA oligonucleotide probe
CC	is useful in diagnostic methods to identify patients having
CC	variant forms of h-erg2 gene, to determine the level of expression
CC	of RNA encoding h-erg2, or to isolate genes homologous to h-erg2 from
CC	other species. The DNA sequence is also useful in gene therapy
CC	techniques to introduce the h-erg2 protein into cells of the target
CC	organs. The present sequence represents the human erg2 protein splice
CC	variant of the invention.
XX	
SO	Sequence 994 AA;
	Query Match 50.0%; Score 3039.5; DB 23; Length 994;
	Best Local Similarity 54.3%; Pred. No. 3.2e-238;
	Matches 670; Conservative 88; Mismatches 152; Indels 325; Gaps 33;
OY	1 MPVRGHVAPONTLEDTIIRFEGOSKRFEITANRVENCAVYICNDGFCELGYSRAEVM 60
DB	1 MPVRGHVAPONTLEDTIIRFEGOSKRFEITANRVENCAVYICNDGFCELGYSRAEVM 60
OY	61 QRPCTCFELHPPRIQRRRAAQAIALGAEKKRYEIAFYRKDGSFCFLTVVVPYKNDG 120
DB	61 QRPCTCFELHPPRIQRRRAAQAIALGAEKKRYEIAFYRKDGSFCFLTVVVPYKNDG 120
OY	121 AVIMEFLINFEVMEKDVWGSPAHDPNHRGPTSMLEPARAKTFRLKLPALKA-LTARESS 179
DB	121 AVIMEFLINFEVMEKDVWGSPAHDPNHRGPTSMLEPARAKTFRLKLPALKA-LTARESS 179
OY	180 VR-----SGGAGAGAGAGVAVVDVLTLPAPASSESLALDEVATMDNHVAGLGAEEERRA 233
DB	180 VR-----SGGAGAGAGAGVAVVDVLTLPAPASSESLALDEVATMDNHVAGLGAEEERRA 233
OY	234 LVFGSGSPRPASRGQLPSPRAHSLNPDASGSSCSLAFTRSRESCASVRRASSADDIEAMRA 293
DB	234 LVFGSGSPRPASRGQLPSPRAHSLNPDASGSSCSLAFTRSRESCASVRRASSADDIEAMRA 293

SQ	Sequence	994	AA;
Query	Match	50.0%;	Score 3039.5; DB 23; Length 994;
Bates	Local Similarity	54.3%;	Pred. No. 3.2e-289;
Matches	670; Conservative	88; Mismatches	152; Indels 325; Gaps 33.
OY	1 MPVRGHAAPONTELDITIRKEFEGOSRKEFIANARVENCAYIYNDGCEGLCYSRRAVM 60		
Dd	1 MPVRGHAAPONTELDITIRKEFEGOSRKEFLIANNQMEECALIIYNDGCEGLFYGRSRYEYM 60		
OY	61 QRPTCDLHGPRTQRRAAQAIAALLGAERKVEIAFYRKDGCFCGLVDVPVKNEGD 120		
Dd	61 QQPCTCDLTPNTPSSAVSRILAOALLGAECCKVDILYYRRKDASFRCGLVDPVPVKNEGD 120		
OY	121 AVIMFIILFEVYMKEWYGSPAHDTNNHGPPTSMLAPORATFRKLPLALKA-LTARESS 179		
Dd	121 AVIMFIILNF-----DLAQLAKCSSRLS 145		
OY	180 VR-----SGGAGGAGACGVVVVDVLTPAAPSSSELALDEVTFAMDNHVAGLPAEEERRA 233		
Dd	146 QRLLSQSLFGSEGSIGRG----- 164		
OY	234 LVFGQSPERPASPGOLPSPFRAHSINLPDASGSSCSLTAFTRSRESCASVRRASSADDIEAMRA 293		

```

Db      165  --GPG-----PG-----
QY      294  GVLPPPRHASTGAMHPLRSLNSTSDSLVRKTIKIPQITLNFVYDLKGPDLASPT 353
Db      170  -----TGR-----GKVTISQIQPFTLNFVFNLEKRRSSST 201
QY      354  SDREIAP-KIKERTHNTEKVTOVLSIGADVLPREYKLOADRIRHWTILHYSPEKAVDW 412
Db      202  TEIETIAPHKRYVERTQNTKERTQVLSIGADVLPREYKLOADRIRHWTILHYSPEKAVDW 261
QY      413  LILLVITYTAFTPYSAFLKETEPECPATECGAACPLAVNDIYIMTVDLILNER 472
Db      262  LILLVITYTAFTPYSAFLSDODESRRGA-CSTTCSPLTVVDIYIMTVDIYINFR 320
QY      473  TTYVANEVEVSHRGRIAVHYFKGWFLLDMVAIPFDLLIFGSGSEE--LIGLKTRARL 529
Db      321  TTYVNTNEVYSHRRIRIAVHYFKGWFLLDMVAIPFDLLIFRTGSEDTTLLIGLKTRARL 380
QY      530  LRLVAVARKLDYSEYGAVALFLMCTFALIAHMLACIWAIGNMEQPHMDSRIGMLHNL 589
Db      381  LRLVAVARKLDYSEYGAVALFLMCTFALIAHMLACIWAIGNVERPYLEHKGMLDSL 440
QY      590  GDOIGKPNSSG-LGSPKIDKYTALYETESSILTSVGFNVSPPTNSEKIFSTCYMLIG 648
Db      441  GVQLGKRTNSDPASGSPVDKYVATLYETSSILTSVGFNVSPPTNSEKIFSTCYMLIG 500
QY      649  SLMTASIFGNVSAIIORLYSGTARYHTQMLRVREFIRPHQIPNPLRQLEEFQHAMSYT 708
Db      501  SLMTASIFGNVSAIIORLYSGTARYHTQMLRVKEIRPHQIPNPLRQLEEFQHAMSYT 560
QY      709  NGIDMNAVLCFPEPCLOADICLHNRSLQHCPRKATKCLRALAMKEFTTHAPGDT 768
Db      561  NGIDMNAVLCFPEPCLOADICLHNRSLQHCPRKATKCLRALAMKEFTTHAPGDT 620
QY      769  LVHAGDLTALYETISRGSEITELRGVYVAIIIGKNDIPEEPNIXVAPKSGVDVALTYC 828
Db      621  LVHAGDLTALYETISRGSEITELRGVYVAIIIGKNDIPEEPNIXVAPKSGVDVALTYC 680
QY      829  DLHRIHRDDLLEVLDMYEFSDHFWSSLEITFNLRDT-----NMIPGSPGSTELEGFSR 883
Db      681  DLHRIHRDDLLEVLDMYEFSDHFWSSLEITFNLRDT-----NMIPGSPGSTELEGFSR 738
QY      884  QRRKRLFRRTDKDTEQPEVSALGPBRAGAPSSRRPGRPGWGESSSGPS--SEPPSS 941
Db      739  -----LSDNQGSGPHE--LGP-----QFSSKGYSLIGPSQ 767
QY      942  EDEGPGRSSSPRLRLVPFSSPPRPGEPGEPLEMEDECKESDTCNPLSAGFVSNIFEFW 1001
Db      768  NSMGAG-----PCAPGHPRDAAPL-----SISDASGLW 795
QY      1002  GDSRGROYQELPCPCPAPTPSLINIPLSSGRRPRGQVY-----SRDALQORINLET 1054
Db      796  PE-----LLQEMP-----PRHSP-QSPQEDPCWPLKLSRLQEQAOQNRLLES 838
QY      1055  RLSDAMATVILQLQROM-----TLVPRATSAVITTPRG-----PIST 1091
Db      839  RVSSDLSSITQLQKPMQGHASYILEAPASNDIALVPIA-SETTSPSPRLPQGFLLPQAQ 897
QY      1092  SP-----LLPVP-----LPTLTL--DSLQVSCPMACEBLPPG-----APELQOE 1130
Db      898  TPSTGDDLDCCPKHRNSSPRMPLAVAMDKTLAPSSF-----QOEDEGLMPLASLAPLE 953
QY      1131  -----GPTTRLSLPGOLGALTQO-PLHRGSDPG 1158
Db      954  VQGLICGFCFS-SLPEHLGVSVPKQLDPQRHGSDDPG 987

```

RESULT 14
 ABG31252
 ID ABG31252 standard; Protein; 958 AA.
 XX AC
 XX ABG31252;

```

DT      21-OCT-2002 (first entry)
XX
DE      Human erg2 (h-erg2) protein variant.
XX
KM      Human; erg2; hypotensive; hypertensive; cytosolic;
KM      antinfertility; nephrotropic; potassium channel inhibitor;
KM      hypotension; hypertension; renal failure; benign prostatic hyperplasia;
KM      prostate cancer; infertility; potassium channel.
XX
OS      Homo sapiens.
XX
PN      WO200242417-A2.
XX
PD      30-MAY-2002.
XX
XX      16-NOV-2001; 2001WO-US43490.
XX      20-NOV-2000; 2000US-249981P.
PR      (MERI ) MERCK & CO INC.
XX
PI      Folander KL, McKenna EJ, Swanson RJ, Liu Y;
XX      WPI; 2002-583376/62.
XX      N-PSDB; ABR88232.
XX
PT      New isolated human-erg2 potassium channel subunit, useful for treatment
PT      of hypertension, hypotension, renal failure, benign prostate
PT      hyperplasia, prostate cancer and infertility
XX
XX      Claim 8; Fig 2B; 53pp; English.
XX
PS      This invention relates to an isolated human erg2 potassium channel
PS      subunit protein. The erg2 protein of the invention is useful for
PS      identifying activators or inhibitors of potassium channels containing
PS      the protein. The erg2 protein is also useful in counter screens for
PS      assays designed to identify activators and inhibitors of other drug
PS      targets. The protein is useful for treating hypotension, hypertension,
PS      renal failure, benign prostatic hyperplasia, prostate cancer, and
PS      infertility. The activators and inhibitors of potassium channels
PS      containing h-erg2 protein, identified using this protein are useful for
PS      treating or preventing conditions as described above, where the
PS      activity of potassium channels containing h-erg2 protein is abnormal.
PS      The nucleic acid encoding the human erg2 protein is useful in
PS      various diagnostic methods, and a DNA or RNA oligonucleotide probe
PS      is useful in diagnostic methods to identify patients having
PS      variant forms of h-erg2 gene, to determine the level of expression
PS      of RNA encoding h-erg2, or to isolate genes homologous to h-erg2 from
PS      other species. The DNA sequence is also useful in gene therapy
PS      techniques to introduce the h-erg2 protein into cells of the target
PS      organs. The present sequence represents the human erg2 protein variant
PS      of the invention.
SQ      Sequence 958 AA:
SQ
Query Match 49.4%; Score 3001.5; DB 23; Length 958;
Best Local Similarity 53.7%; Pred. No. 3.8e-235;
Matches 659; Conservative 83; Mismatches 139; Indels 347; Gaps 30;
QY      1 MPVRGHAAPONTFLDITIRKEGSRKFTIANRVECAVITYCDGFCGLGSRAYM 60
Db      1 MPVRGHAAPONTYLTDTIRKEGSRKFTIANRVECAVITYCDGFCGLGSRAYM 60
QY      61 QRPCTCDLHPGRQORRAAOIAQALGAERKVEIARYRKGSCFLVYVVPYKNDG 120
Db      61 QQPCTCDLTPGPNTPSSAVSKLAQALGAECCKVDILYRRDASSFCRLVYVVPYKNDG 120
QY      121 AVIMEFLNFVEYMEKDWGSPAHDTHRGPTSWLAPGRAFTFLKLPALIA-LTARESS 179
Db      121 AVIMEFLNF-----DLAQLAKCSSRSLIS 145
QY      180 VR-----SGAGAGAGATVVDVLTLPAPSSSLALDEVYTMADNHVAGLGRAEERRA 233

```

Db	146	QRLSQSLGSEGSIGRGC-----	164
OY	234	LVPGSPSPRAPGOLPSPRANSLNDASGSSCLARTSRSCASVRRASSADIEAMRA	293
Db	165	--GRC-----PG-----	169
OY	294	GVLPPPPHNAHSTGANHPLRSLGLNSTSDSLVRYTITKIPOTITLNFVDLKCDPLASPT	353
Db	170	-----TGR-----GKYRTISOIPQFTLNFVEFNLEKRRSSST	201
OY	354	SDREITAP-KIKERTHNTEKTVOLSLGADVLPEYKLOAPRIHMTLTHYSPFAWVM	412
Db	202	TEIELAPHKVEYERQONTTEKTVOLSLGADVLPEYKLOAPRIHMTLTHYSPFAWVM	261
OY	413	LILLVITYAVFTPYSAAFLLKETEECPATECGACOPLAVADVLIVIMFIVDLINER	472
Db	262	LILLVITYAVFTPYSAAFLLSDQESRGA-CSYTSGLPYVDLIVIMFIVDLINER	320
OY	473	TTYVANEVYVSHPCRIVNHYFKGMFLIDMAAIFPDLIIEGSGSEB--LIGLTKTARL	529
Db	321	TTYVNTNEVYVSHPRIRIVNHYFKGMFLIDMAAIFPDLIIERTGSDETTLIGLTKTARL	380
OY	530	LRIYVAVKLDYRYSYGAAYFLMCTFALLAHMLACTIYXAGNMEOHMDSRIGIMLNL	589
Db	381	LRIYVAVKLDYRYSYGAAYFLMCTFALLAHMLACTIYXAGNERPRLERHKIGIMLNL	440
OY	590	GDQICKPYNSSG-LGGSPIDKRYVALYFTFESSLSYVGNGVSPMTNSEKIPISICVMLIG	648
Db	441	GVOLGKRRNGSDPAGSPSVODKYVALYFTFESSLSYVGNGVSPMTNSEKIVSICVMLIG	500
OY	649	SLMTASIFGNVSAILQRLYSGTARYHTOMLRVREIRFHQIPIPNLRLEEYFOHAWSYT	708
Db	501	SLMTASIFGNVSAILQRLYSGTARYHTOMLRVREIRFHQIPIPNLRLEEYFOHAWSYT	560
OY	709	NGIDMNAVKEPPELOADICLHNLRLDQCKPFRGATKGLRYLAKMFKTTNAPRGDT	768
Db	561	NGIDMNAVKEPPELOADICLHNLRLDQCKPFRGATKGLRYLAKMFKTTNAPRGDT	620
OY	769	LVAHGDLTALYFISRGSEIILRGVYVAIILGKNPIFGPBLUYARPGKSNQDVALTYC	828
Db	621	LVAHGDLTALYFISRGSEIILRDYVAIILGKNPIFGPBPVSLHAPGKSSADVALTYC	680
OY	829	DLHKIHRDDELVLDMYPEFSDFHFWSSLEITFNLDT-----NMLPGSPGTELECGFSR	883
Db	681	DLHKIQRADLLEVLDMYPAFAESFWSKLEVFENLMDAAGLHSPRQAPGSDHQGF--	738
OY	884	QRKRLSRRTKDKDTEOPGEVYALGPRGAGPSSRCRPGGPMGESPPSGSPSPESSED	943
Db	739	-----LSDNOSD-----AAPRLSIDASGLM-----PELLOE	765
OY	944	EGPGRSSPRLRLVPPSPRPGEPGSRPLMBDCSSDTCNPLSGAFSGVNTISPFMD	1001
Db	766	-----MPRHSPQSPQEDB-----DC-----W--	782
OY	1004	SRGROYELPRCOPATPSELINIPLSSPGRPRGVDERSRLDALORLNRLETSLADMATV	1063
Db	783	-----PLK-----LGSNLEQLOAQAMNRLRESVSDLSKRI	811
OY	1064	LQLOLQROM-----TLVPRAYSAVTTPGP-----GPTST-----	1093
Db	812	LQLOLQPMPOGHASYILREPASNDIALPIA-SETTSGPRLPGFLLPPAPGYSGLDD	870
OY	1092	-----SPLLPVSPRLPTLLDLSLSQVSOFCACELPRG-----ABELPOE-----G	1133
Db	871	CSPKHNSSPKMP--HLAVATDKTILAPSE-----QOEPGIMPLPLASPLRLEVGLIG	924
OY	1132	PTRRLSLPGOLGALTSQ--PLRHGSDPG	1158
Db	925	PCFS-SLEPHIGSVKQOLDPQRHSGDPG	951

RESULT 15
ABP72167

ID	ABP72167	standard; Protein; 958 AA.
XX		
AC	ABP72167;	
XX		
DT	22-APR-2003	(first entry)
XX		
DE	Human potassium channel 52908.	
XX		
KW	Human; potassium channel; 52908; nootropic; neuroprotective; analgesic; cardiovascular; cardiac; cytosolic; immunosuppressive; anti-inflammatory; osteopathic; hepatotropic; virucide; gene therapy.	
KW		
XX		
OS	Homo sapiens.	
XX		
EH	Key	Location/Qualifiers
FT	Domain	41..59
FT		/note= "PAS domain"
FT	Domain	93..135
FT		/note= "PAC domain"
FT	Domain	261..282
FT		/note= "Transmembrane domain"
FT	Domain	295..317
FT		/note= "Transmembrane domain"
FT	Domain	345..361
FT		/note= "Transmembrane domain"
FT	Domain	371..390
FT		/note= "Transmembrane domain"
FT	Domain	398..420
FT		/note= "Transmembrane domain"
FT	Domain	491..515
FT		/note= "Transmembrane domain"
FT	Domain	341..580
FT		/note= "Cyclic nucleotide gated channel transmembrane region"
FT	Domain	463..482
FT		/note= "Pore Loop domain"
FT	Region	478..480
FT		/note= "Potassium channel pore motif"
FT	Region	608..699
FT		/note= "Cyclic nucleotide gated channel transmembrane region"
FT		
PN	WO2003006679-A2.	
XX		
PD	23-JAN-2003.	
XX		
PF	10-JUL-2002; 2002MO-US22456.	
XX		
PR	10-JUL-2001; 2001US-304243P.	
XX		
PA	(MILL-) MILLENNIUM PHARM INC.	
XX		
PI	Curtis RAJ;	
XX		
DR	WPI: 2003-221759/21.	
XX	N-PsDB: ABZ58129.	
XX		
PT	New potassium channel family member, designated 52908, nucleic acids and proteins, useful for diagnosing, preventing, alleviating or treating e.g. pain and metabolic disorders, cardiovascular disorders or viral diseases	
PT	-	
XX		
PS	Claim 13; Page 100-102; 118pp; English.	
XX		
CC	The present sequence is the protein sequence of human 52908, a protein having the structural characteristics of a member of the potassium channel family. The invention provides 52908 nucleic acid molecules, recombinant expression vectors, host cells, transgenic animals in which a 52908 gene has been introduced or disrupted, 52098 proteins, fusion proteins, antigenic peptides and anti-52098 antibodies. The 52908 nucleic acids and proteins are useful in screening assays, predictive medicine (e.g. diagnostic	

CC and prognostic assays, monitoring clinical trials, and
 CC pharmacogenomics), and in methods of treatment of central nervous
 CC system (e.g. neurological disorders), pain and metabolic disorders,
 CC cardiovascular disorders, cellular proliferative, growth,
 CC differentiation and/or migration disorders, immune e.g.
 CC inflammatory disorders, disorders associated with bone metabolism,
 CC endothelial cell disorders, liver disorders, and viral diseases.

XX Sequence 958 AA:

Query Match 49.4%; Score 3001.5; DB 24; Length 958;

Best Local Similarity 53.7%; Pred. No. 3.8e-235; Mismatches 139; Indels 347; Gaps 30;

Matches 659; Conservative 83; Pident 139; Indels 347; Gaps 30;

QY 1 MPVRGHVAPONTFLDTIRKFEQSRKFIITANAVENCAYICNDGFCGLGYSRAYM 60
 DB 1 MPVRGHVAPONTFLDTIRKFEQSRKFIITANAVENCAYICNDGFCGLGYSRAYM 60
 QY 61 QRPCTDFLHGPFRORAAQIAQALGAERKVEIAFYRKDGSCFLCLVDVVPYKNDG 120
 DB 61 QRPCTDFLHGPFRORAAQIAQALGAERKVEIAFYRKDGSCFLCLVDVVPYKNDG 120
 QY 121 AVIMFLNEFYVMEDMGWSPAHDNHNHGPPTSWLAPRAKTRFKLKLALLA-LTARESS 179
 DB 121 AVIMFLNEFYVMEDMGWSPAHDNHNHGPPTSWLAPRAKTRFKLKLALLA-LTARESS 179
 QY 180 VR-----SGGAGAGAPGAVVVDVLTAPAASSESLADEVTAMDNHVGIGPAEERRA 233
 DB 146 QRLLSQSFSGSESHGRPG----- 164
 QY 234 LVGCGSPPRSAPGQLPSRAHSLNDASGSSCLARTSRSCASVRRASSADIEAMRA 293
 DB 165 --GPG-----PG----- 169
 QY 294 GVLPPRRHASTGAMHPLRSGLINSTSDSLVRYRTISKIPQITLNFYDLKGPFLASPT 353
 DB 170 -----TGR-----GKRTTISQIPQITLNFYDLKGPFLASPT 201
 QY 354 SDRELIAP-KIKERTHNTKVTQVLSGADVLPKYKLAAPRIHRTILHYSFKAVMDW 412
 DB 202 TEELIAPHKVERQNTKVTQVLSGADVLPKYKLAAPRIHRTILHYSFKAVMDW 261
 QY 413 LILLIYITAVFTPYSAFLKETEBCGATACOPLAVIDLIDIMEFVIDILNFR 472
 DB 262 LILLIYITAVFTPYSAFLKETEBCGATACOPLAVIDLIDIMEFVIDILNFR 320
 QY 473 TTYVANEVYSHGRIVAHYFKGWFLLDMVAALPEDLIFRTGSDETTLGLKTARL 529
 DB 321 TTYVANEVYSHGRIVAHYFKGWFLLDMVAALPEDLIFRTGSDETTLGLKTARL 380
 QY 530 LRLVVARKLDRYSEYGAAYFLMCTPALLAHMLACIMYAIGNNEQPHMDSRIGMLNL 589
 DB 381 LRLVVARKLDRYSEYGAAYFLMCTPALLAHMLACIMYAIGNNEQPHMDSRIGMLNL 440
 QY 590 GDQIGKPYNSG-LGGSIKDKYVATLYFTSSSLTSGFGVNSPNTNSEKIFSTCVMLIG 648
 DB 441 GVOIGKRYNSGDPASGPEVDKYVATLYFTSSSLTSGFGVNSPNTNSEKIFSTCVMLIG 500
 QY 649 SLMTASTIGNSATIIRLYSGTARHTOMLRYREFIRFHOJPNPLRQLEEFQHAWSYT 708
 DB 501 SLMTASTIGNSATIIRLYSGTARHTOMLRYREFIRFHOJPNPLRQLEEFQHAWSYT 560
 QY 709 NGIDMNAVKGFPEDCLQADICLHLNRSLLQCHKPFGATKCLRALAMKFTTHAPPGDT 768
 DB 561 NGIDMNAVKGFPEDCLQADICLHLNRSLLQCHKPFGATKCLRALAMKFTTHAPPGDT 620
 QY 769 LVHAGDLITALLYFISRGSEILRDGVVVALLGKNDIFGEPLNLXARPGKSGDVARALTYC 828
 DB 621 LVHAGDLITALLYFISRGSEILRDGVVVALLGKNDIFGEPLNLXARPGKSGDVARALTYC 680
 QY 829 DLHKIHRDLEVLDMYPSFDHFWSSLEITFNLRDT-----NMTPSGSPSTEEGFSR 883
 DB 681 DLHKIHRDLEVLDMYPSFDHFWSSLEITFNLRDT-----NMTPSGSPSTEEGFSR 738

QY 884 QRRKLSFRRTDXDEOPGEVSALGPRAGAPSSRGRPGWPGESSPSGSPRESSED 943
 DB 739 -----LSDNGSD-----AAPPLISDASGLM-----PELLOE 765
 QY 944 EGPGRSSPRLVPPSSRPPEGEPPELMEDECKSSDTCNPLSGAFSGVSNITFSWGD 1003
 DB 766 -----MPRHSPOSQEDP-----DC-----W-- 782
 QY 1004 SRGRQYELPPOCAPTPSLNTPSSPGRPRGDSRDLALQRLNLETRLSADMATV 1063
 DB 783 -----PLK-----LGSRLQLOQAMNRLSRYSSDLSRT 811
 QY 1064 LQLQROM-----TLVPPAYSAVTPGP-----GPTSP----- 1091
 DB 812 LQLQROMPGHASYTLBPASNDLALVPIA-SETTSPGRLPGFLPPAQTPSYGDLDD 870
 QY 1092 -----SPLPSPPLFTLDSLSOVSQFMACELEPPG-----APELPOE-----G 1131
 DB 871 CSPKHNSSPRMP--HLAVATDKTLAPSE-----QOEPGGLMPPPLASPLHPLVQGLIG 924
 QY 1132 PTRRLSLPGOLGALTSQ--PLHRHGSDPG 1158
 DB 925 PCFS-SLPEHLGSVPKQLDFQRHGSDPG 951

Search completed: September 23, 2003, 17:37:01
 Job time : 81 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 23, 2003, 11:02:09 ; Search time 34 Seconds

(without alignments)
1603.057 Million cell updates/sec

Title: US-10-000-151b-3

Perfect score: 6079
Sequence: 1 MPVRRGHVAPONTFLDITR.....GOLGALTQPLRRHSDPGS 1159.

Scoring table: BLOSUM62
Gapop 10.0 , Gapept 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	6079	100.0	1159	1 KCH2_HUMAN	Q12809 homo sapien
2	5942	97.7	1161	1 KCH2_RABIT	Q8WNY2 oryctolagus
3	5872.5	96.6	1158	1 KCH2_CANFA	Q9TSZ3 canis famli
4	5797.5	95.4	1162	1 KCH2_MOUSE	Q35219 mus musculu
5	5790	95.2	1163	1 KCH2_RAT	Q08962 rattus norv
6	3312.5	54.5	1195	1 KCH7_RAT	Q54852 rattus norv
7	3283.5	54.0	1195	1 KCH7_MOUSE	Q9ER47 mus musculu
8	3255	53.5	1196	1 KCH7_HUMAN	Q9NS40 homo sapien
9	3043.5	50.1	994	1 KCH6_HUMAN	Q9H252 homo sapien
10	2982	49.1	950	1 KCH6_RAT	Q54853 rattus norv
11	2275.5	37.4	526	1 KCH2_CHICK	Q9PT84 gallus gall
12	1636	26.9	1017	1 KCH4_HUMAN	Q9UG05 homo sapien
13	1616	26.6	1017	1 KCH4_RAT	Q9RT15 rattus norv
14	1568	25.8	1107	1 KCH8_HUMAN	Q96142 homo sapien
15	1551	25.5	1102	1 KCH8_RAT	Q9QWS8 rattus norv
16	1540.5	25.3	1083	1 KCH3_HUMAN	Q9ULI8 homo sapien
17	1537.5	25.3	1087	1 KCH3_RAT	Q89047 rattus norv
18	1527.5	25.1	1087	1 KCH3_MOUSE	Q9WVJ0 mus musculu
19	1448.5	23.8	1174	1 CIKE_DROME	Q02280 drosophila
20	1409.5	23.2	962	1 KCH1_RAT	Q63472 rattus norv
21	1406.5	23.1	987	1 KCH1_BOVIN	Q18965 bos taurus
22	1402.5	23.1	989	1 KCH1_HUMAN	Q95259 homo sapien
23	1402	23.1	989	1 KCH1_MOUSE	Q60603 mus musculu
24	1399.5	23.0	988	1 KCH5_HUMAN	Q8NCM2 homo sapien
25	1398.5	23.0	988	1 KCH5_RAT	Q9EP19 rattus norv
26	1193	19.6	876	1 KCH8_MOUSE	P93111 mus musculu
27	804	13.2	162	1 KCH4_CAVPO	Q08703 cavia porce
28	595.5	9.8	1175	1 KCH4_RABIT	Q9TV66 oryctolagus
29	581	9.6	1186	1 KCH4_RAT	Q9JKA7 rattus norv
30	578	9.5	1186	1 KCH4_MOUSE	Q9Y3G4 mus musculu
31	563	9.3	1203	1 KCH4_HUMAN	Q9UL51 homo sapien
32	551.5	9.1	889	1 KCH2_HUMAN	Q9UL51 homo sapien
33	543	8.9	863	1 HCN2_MOUSE	Q88703 mus musculu

34	529.5	8.7	822	1 HCN1_RABIT	Q8WZS1 oryctolagus
35	527.5	8.7	890	1 HCN1_HUMAN	Q60741 homo sapien
36	524.5	8.6	834	1 HCN2_RAT	Q9JKA9 rattus norv
37	515.5	8.5	910	1 HCN1_MOUSE	Q88704 mus musculu
38	512.5	8.4	774	1 HCN3_HUMAN	Q9P123 homo sapien
39	509.5	8.4	780	1 HCN3_RAT	Q9JKA8 rattus norv
40	507.5	8.3	910	1 HCN1_RAT	Q9JKB0 rattus norv
41	506.5	8.3	779	1 HCN3_MOUSE	Q88705 mus musculu
42	476	7.8	97	1 KCH2_PIG	Q9ULI4 sus scrofa
43	451.5	7.4	665	1 CNG3_DROME	Q24278 drosophila
44	440	7.2	706	1 CNG3_BOVIN	Q29441 bos taurus
45	438	7.2	694	1 CNG3_HUMAN	Q16281 homo sapien

ALIGNMENTS

RESULT 1
KCH2_HUMAN STANDARD; PRT; 1159 AA.
AC Q12809: 075418: 075680: Q9PT72; Q9BUT7; Q9H3P0:
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Potassium voltage-gated channel subfamily H member 2 (Ether-a-go-go
DE related gene potassium channel 1) (H-ERG) (ERG1) (Ether-a-go-go
DE related protein 1) (Eag related protein 1) (eag homolog).
GN KCHH2 OR HERG OR HERG1 OR ERG OR ERG1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
NCBI_Taxid=9606;
[1]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RC TISSUE-Hippocampus;
RX MEDLINE=94211879; PubMed=8159766;
RA Warmke J.W., Ganetzky B.;
RT "A family of potassium channel genes related to eag in Drosophila and
RT mammals".
RL Proc. Natl. Acad. Sci. U.S.A. 91:3438-3442(1994).
[2]
RP SEQUENCE FROM N.A. (ISOFORM 1), AND VARIANT LQ72 CYS-534.
RX MEDLINE=98260867; PubMed=9600240;
RA Itoh T., Tanaka T., Nagai R., Kamiya T., Sawayama T., Nakayama T.,
RA Tomoike H., Sakurada H., Yazaki Y., Nakamura Y.;
RT "Genomic organization and mutational analysis of HERG, a gene
RT responsible for familial long QT syndrome".
RL Hum. Genet. 102:435-439(1998).
[3]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RA Donnie D., Chapman C.G., Punia P., Rice S., Bahmani F., Mirdock P.,
RA Pearson N., Randall A.D., Meadows H.J.;
RT "Potent inhibition of HERG K⁺ channels by the neuroprotective agent
RT Sildenafil".
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
[4]
RP SEQUENCE FROM N.A. (ISOFORM 3).
RC TISSUE-Heart;
RX MEDLINE=21269186; PubMed=11374908;
RA Soejima H., Kawamoto S., Akai J., Miyoshi O., Arai Y., Morohka T.,
RA Matsuo S., Nishikawa N., Kimura A., Okubo K., Mukai T.;
RT "Isolation of novel heart-specific genes using BodyMap database".
RL Genomics 74:115-120(2001).
[5]
RP SEQUENCE FROM N.A. (ISOFORM 2).
RA Cioconini O., Guastri L., Balzi M., Becchetti A., Wanke E., Olivetto M.,
RA Wyomere R.S., Arcangeli A.;
RT "Cell cycle-dependent expression of HERG1 and HERG1b isoforms in tumor
RT cells".
RL Submitted (OCT-2002) to the EMBL/GenBank/DBJ databases.
[6]
RP SEQUENCE OF 1-376 FROM N.A. (ISOFORM 2).
RC TISSUE-Heart;

RX MEDLINE-98012815; PubMed-9351462;
RA London B., Trudeau M.C., Newton K.P., Beyer A.K., Copeland N.G.,
RA Gilbert D.J., Jenkun N.A., Satler C.A., Robertson G.A.;
RT "Two isoforms of the mouse ether-a-go-go-related gene coassemble to
RT form channels with properties similar to the rapidly activating
RT component of the cardiac delayed rectifier K⁺ current.";
RL Circ. Res. 81:870-878(1997).
RN [17]
RP SEQUENCE OF 1-376 FROM N.A. (ISOFORM 2).
RC TISSUE=Heart atrium;
RX MEDLINE-98012799; PubMed-9351446;
RA Lees-Miller J.P., Kondo C., Wang L., Duff H.J.;
RT "Electrophysiological characterization of an alternatively processed
RT ERG K⁺ channel in mouse and human hearts.";
RL Circ. Res. 81:719-726(1997).
RN [18]
RP SEQUENCE OF 27-1159 FROM N.A. (ISOFORM 1).
RA Yang P., Paulussen A., Verhasselt P., Crabbe R., Luyten W.,
RA Armstrong M.;
RT "Analysis of the human HERG gene: Intron localisation and
RT identification of a novel inherited mutation associated with long
RT QT.";
RL Submitted (Aug-1998) to the EMBL/GenBank/DBS databases.
RN [19]
RP SEQUENCE OF 795-1159 FROM N.A. (ISOFORM 3).
RC TISSUE=Heart ventricle;
RX MEDLINE-98438490; PubMed-9765245;
RA Kupersmidt S., Snyder D.J., Raes A., Roden D.M.;
RT "A K⁺ channel splice variant common in human heart lacks a C-terminal
RT domain required for expression of rapidly activating delayed
RT rectifier current.";
RL J. Biol. Chem. 273:27231-27235(1998).
RN [110]
RP SEQUENCE OF 59-1159 FROM N.A. (ISOFORM 4), AND SEQUENCE OF 711-1159
RP FROM N.A. (ISOFORM 1/2).
RC TISSUE=Brain;
RX MEDLINE-22388257; PubMed-12477932;
RA Strussberg R.L., Felting E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shemen C.M., Schuler G.D.,
RA Altshul S.F., Zeeberg B., Buelow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Staelen M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Rana S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosok S.A., McMan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butlerfield Y.S.N., Krzywinski M.I., Skalska U., Smalhus D.E.,
RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [111]
RP MUTAGENESIS OF ASN-598; ASN-629 AND SER-631, AND N-GLYCOSYLATION.
RX MEDLINE-22057172; PubMed-12063277;
RA Gong Q., Anderson C.L., January C.T., Zhou Z.;
RT "Role of glycosylation in cell surface expression and stability of
RT HERG potassium channels.";
RL Am. J. Physiol. 283:H77-H84(2002).
RN [112]
RP MUTAGENESIS OF SER-283; SER-890; THR-895 AND SER-1137, AND
RP PHOSPHORYLATION.
RX MEDLINE-20299343; PubMed-10837251;
RA Cui J., Melman Y., Palma E., Fishman G.I., McDonald T.V.;
RT "Cyclic AMP regulates the HERG K(+) channel by dual pathways.";
RL Curr. Biol. 10:671-674(2000).
RN [113]
RP INTERACTION WITH KCNE1.

RX MEDLINE-97373956; PubMed-9230439;
RA McDonald T.V., Yu Z., Ming Z., Palma E., Meyers M.B., Wang K.-W.,
RA Goldstein S.A., Fishman G.I.;
RT "A link-HERG complex regulates the cardiac potassium current I(Kr).";
RL Nature 388:289-292(1997).
RN [114]
RP INTERACTION WITH KCNE2.
RX MEDLINE-99235979; PubMed-10219239;
RA Abbott G.W., Sesti F., Splawski I., Buck M.E., Lehmann M.H.,
RA Timothy K.W., Keating M.T., Goldstein S.A.;
RT "MiRP1 forms IKr potassium channels with HERG and is associated with
RT cardiac arrhythmia";
RL Cell 97:175-187(1999).
RN [115]
RP X-RAY CRYSTALLOGRAPHY (2.6 ANGSTROMS) OF 1-135, AND MUTAGENESIS OF
RP PHE-29 AND TYR-43.
RX MEDLINE-99059500; PubMed-9845367;
RA Morais Cabral J.H., Lee A., Cohen S.L., Chait B.T., Li M.,
RA Mackinnon R.;
RT "Crystal structure and functional analysis of the HERG potassium
RT channel N terminus: a eukaryotic PAS domain.";
RL Cell 95:649-655(1998).
RN [116]
RP VARIANTS LOT2 ASP-470; VAL-561 AND SER-628.
RX MEDLINE-95196272; PubMed-7889573;
RA Curran M.E., Splawski I., Timothy K.W., Vincent G.M., Green E.D.,
RA Keating M.T.;
RT "A molecular basis for cardiac arrhythmia: HERG mutations cause long
RT QT syndrome.";
RL Cell 80:795-803(1995).
RN [117]
RP VARIANT LOT2 MET-822.
RX MEDLINE-97071892; PubMed-8914737;
RA Satler C.A., Walsh E.P., Vesely M.R., Plummer M.H., Ginsburg G.S.,
RA Jacob H.J.;
RT "Novel missense mutation in the cyclic nucleotide-binding domain of
RT HERG causes long QT syndrome.";
RL Am. J. Med. Genet. 65:27-35(1996).
RN [118]
RP VARIANT LOT2 ARG-593.
RX MEDLINE-96259954; PubMed-8635257;
RA Benson D.W., Macrae C.A., Vesely M.R., Walsh E.P., Seidman J.G.,
RA Seidman C.E., Satler C.A.;
RT "Missense mutation in the pore region of HERG causes familial long QT
RT syndrome.";
RL Circulation 93:1791-1795(1996).
RN [119]
RP VARIANT LOT2 THR-561.
RX MEDLINE-97031865; PubMed-8877771;
RA Dausse E., Berthet M., Denjoy I., Andre-Fouet X., Craud C.,
RA Benacour M., Faure S., Coumel P., Schwartz K., Guicheney P.;
RT "A mutation in HERG associated with notched T waves in long QT
RT syndrome";
RL J. Mol. Cell. Cardiol. 28:1609-1615(1996).
RN [120]
RP VARIANTS LOT2 ILE-474; HIS-611; VAL-614 AND LEU-630.
RX MEDLINE-97176600; PubMed-9024139;
RA Tanaka T., Nagai R., Tomoike H., Takata S., Yano K., Yabuta K.,
RA Haneda N., Nakano O., Shibata A., Sawayama T., Kasai H., Yazaki Y.,
RA Nakamura Y.;
RT "Four novel KVLQT1 and four novel HERG mutations in familial long-QT
RT syndrome";
RL Circulation 95:565-567(1997).
RN [121]
RP VARIANTS LOT2 CYS-572; ASP-588; VAL-614 AND ALA-630.
RX MEDLINE-98360095; PubMed-9693036;
RA Splawski I., Shen J., Timothy K.W., Vincent G.M., Lehmann M.H.,
RA Keating M.T.;
RT "Genomic structure of three long QT syndrome genes: KVLQT1, HERG, and
RT KCNE1";
RL Genomics 51:86-97(1998).
RN [122]
RP VARIANTS LOT2 LEU-612; VAL-614; ASP-629; SER-629 AND SER-633.

Query Match 100.0%; Score 6079; DB 1; Length 1159;
 Best Local Similarity 100.0%; Pred. No. 3.3e-307;
 Matches 1159; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 MPVRGHVAPQNTFLDIIIRKFEQOSRKFTIANARVENCAYIYCNDFCELCGYSRAEVM 60
DB 1 MPVRGHVAPQNTFLDIIIRKFEQOSRKFTIANARVENCAYIYCNDFCELCGYSRAEVM 60
QY 61 QRECTDDELHGPRTORAAQIAOALLGAERKEVEIAFYRKDSCFCLVDVVPVKNEDG 120
DB 61 QRECTDDELHGPRTORAAQIAOALLGAERKEVEIAFYRKDSCFCLVDVVPVKNEDG 120
QY 121 AVIMFLINFEVMEKDMVSGSPADTNHGPPTSMALPGRKATFRKLKLPALLATARESSV 180
DB 121 AVIMFLINFEVMEKDMVSGSPADTNHGPPTSMALPGRKATFRKLKLPALLATARESSV 180
QY 181 RSGGAGAGAGAVVVDVLTTPAAPSSSEIALDEVTAMDNVAGLGAEBERRALVGPSP 240
DB 181 RSGGAGAGAGAVVVDVLTTPAAPSSSEIALDEVTAMDNVAGLGAEBERRALVGPSP 240
QY 241 PRAPQOLPSPRAHSLNPDASGSSCLARTRSRSCASVRRASSADDIEMRAGVLP 300
DB 241 PRAPQOLPSPRAHSLNPDASGSSCLARTRSRSCASVRRASSADDIEMRAGVLP 300
QY 301 RHASTGAMHPLRGLNSTSDSLVRYRTISKIPQTLNFPVDLKGDFELASPTSDREIIA 360
DB 301 RHASTGAMHPLRGLNSTSDSLVRYRTISKIPQTLNFPVDLKGDFELASPTSDREIIA 360
QY 361 PKIKERTHNTEKVTQVLSGADVLPEYKLOAPRIHRMTILHYSPEKAVMDMLLLVIY 420
DB 361 PKIKERTHNTEKVTQVLSGADVLPEYKLOAPRIHRMTILHYSPEKAVMDMLLLVIY 420
QY 421 TAVFTYSAFLKETEPEGPATCEGACOPLAVDLIYIMFTVDLINFRTTYVANE 480
DB 421 TAVFTYSAFLKETEPEGPATCEGACOPLAVDLIYIMFTVDLINFRTTYVANE 480
QY 481 EVVSHFGRIAVHYFKGWFLLDMYAAIPEFDLIFGSGSEELIGLKTARLLRVAVARKLD 540
DB 481 EVVSHFGRIAVHYFKGWFLLDMYAAIPEFDLIFGSGSEELIGLKTARLLRVAVARKLD 540
QY 541 RSEYGAANVELLMCTFPALIAHMLACTIWAIGNMOPHMDSRIGMLHNLDOIGKPYNS 600
DB 541 RSEYGAANVELLMCTFPALIAHMLACTIWAIGNMOPHMDSRIGMLHNLDOIGKPYNS 600
QY 601 GIGGPRSTKDKYVVALYTFSSLSYVGVNVSPTNNSKIFSIYMLIGSLMYASIFGVNS 660
DB 601 GIGGPRSTKDKYVVALYTFSSLSYVGVNVSPTNNSKIFSIYMLIGSLMYASIFGVNS 660
QY 661 AIIORLYSGTARYHTQMLRVREFIRFHQIENPLRORLEEFYFOHAMSTYNGIDMNAVYKGF 720
DB 661 AIIORLYSGTARYHTQMLRVREFIRFHQIENPLRORLEEFYFOHAMSTYNGIDMNAVYKGF 720
QY 721 PECLQADICLHLNRSLLQHCPRGATKGLRALAMKFTTHAPGPDVTHAGDULTALY 780
DB 721 PECLQADICLHLNRSLLQHCPRGATKGLRALAMKFTTHAPGPDVTHAGDULTALY 780
QY 781 F1RSGSTEILRGVVAAILGKNIDIFGEPMLYARPGSNDVARYLTCDLHKIHRDLE 840
DB 781 F1RSGSTEILRGVVAAILGKNIDIFGEPMLYARPGSNDVARYLTCDLHKIHRDLE 840
QY 841 VLMYPEFSDHFWSSLETFNLDNTNIPGSPSTELGEGFSQORRKKLSFRRTDQTE 900
DB 841 VLMYPEFSDHFWSSLETFNLDNTNIPGSPSTELGEGFSQORRKKLSFRRTDQTE 900
QY 901 QPEVSALGPRAGAGSPSRGPGWGESPSGSSPESSEDEGPGRSSPLRLVFPSS 960
DB 901 QPEVSALGPRAGAGSPSRGPGWGESPSGSSPESSEDEGPGRSSPLRLVFPSS 960
QY 961 PRPPGPRGEPPLMECEKSSDTCNPLSGAFSGVSNIFSWGDSRGQYDELPRCAPTP 1020
DB 961 PRPPGPRGEPPLMECEKSSDTCNPLSGAFSGVSNIFSWGDSRGQYDELPRCAPTP 1020

```

```

QY 1021 SLNINPLSSPGRRRGVDESRLDALQRLNRLFRLSADMAVYLQLLQRMVLPVAYSA 1080
DB 1021 SLNINPLSSPGRRRGVDESRLDALQRLNRLFRLSADMAVYLQLLQRMVLPVAYSA 1080
QY 1081 VTTPEGPSTSPLLPVPSPPTLTLDSLVSQFMACEELPPGAPELPQSGPTRLSLPG 1140
DB 1081 VTTPEGPSTSPLLPVPSPPTLTLDSLVSQFMACEELPPGAPELPQSGPTRLSLPG 1140
QY 1141 QLGALTSOPLHRHSGDPGS 1159
DB 1141 QLGALTSOPLHRHSGDPGS 1159

RESULT 2
KCH2_RABIT
ID Q8WNV2: 002731; 019119; 097586; 09Y06;
AC 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DE 28-FEB-2003 (Rel. 41, Last annotation update)
DE Potassium voltage-gated channel subfamily H member 2 (Ether-a-go-go related gene potassium channel 1) (ERG1) (RERG) (ra-ery) (Ether-a-go-go related protein 1) (Eag related protein 1).
GN KCHN2 OR ERG.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
RA Wlitcher H.J., Hancock J.C., Levi A.J., Meech R.W.;
RT "RERG - rabbit ventricular ERG K+ channel subunit."
RL Submitted (JAN-1997) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE OF 411-571 FROM N.A. (ISOFORM 1/2).
RX MEDLINE=97164986; PubMed=9012748;
RA Wymore R.S., Gintant G.A., Wymore R.T., Dixon J.E., McKinnon D., Cohen I.S.;
RT "Issue and species distribution of mRNA for the IKr-like K+ channel, RT erg."
RL Circ. Res. 80:261-268(1997).
CC -1- FUNCTION: Pore-forming (alpha) subunit of voltage-gated inwardly rectifying potassium channel. Channel properties are modulated by cAMP and subunit assembly. Mediates the rapidly activating component of the delayed rectifying potassium current in heart (IKr) (by similarity).
CC -1- SUBUNIT: The potassium channel is probably composed of a homo- or heterotetrameric complex of pore-forming alpha subunits that can associate with modulating beta subunits. Heteromultimer with KCNH6/ERG2, KCNH7/ERG3, KCNE1 and KCNE2 (by similarity).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- ALTERNATIVE PRODUCTS:
CC Event-Alternative splicing; Named isoforms=2;
CC Name=1;
CC IsoId=Q8WNV2-1; Sequence=Displayed;
CC Name=2;
CC IsoId=Q8WNV2-2; Sequence=VSP_000971;
CC -1- TISSUE SPECIFICITY: Detected in heart, both in atrium and in left ventricle.
CC -1- DOMAIN: The segment 54 is probably the voltage-sensor and is characterized by a series of positively charged amino acids at every third position.
CC -1- PM: Phosphorylated on serine and threonine residues (by similarity).
CC -1- SIMILARITY: Belongs to the potassium channel family. H (Eag) subfamily.
CC -1- SIMILARITY: Contains 1 cyclic nucleotide-binding domain.
CC -1- SIMILARITY: Contains 1 PAS (PER-ARNT-SIM) dimerization domain.
CC -1- SIMILARITY: Contains 1 PAS-associated C-terminal (PAC) domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its

```

CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: U87513; AAC86812.1; -
 DR EMBL: AF068736; AAC98425.1; -
 DR EMBL: AF05061; AAD39357.1; -
 DR EMBL: U75212; AAC48723.1; -
 DR InterPro: IPR000595; CNMP_binding.
 DR InterPro: IPR003967; Erg_channel.
 DR InterPro: IPR005821; Ion_trans.
 DR InterPro: IPR001622; K+channel_pore.
 DR InterPro: IPR005820; M+channel_nig.
 DR InterPro: IPR001610; PAC.
 DR InterPro: IPR000700; PAS-assoc_C.
 DR InterPro: IPR000014; PAS_domain.
 DR Pfam: PF00027; CNMP_binding.1.
 DR Pfam: PF00520; ion_trans.1.
 DR Pfam: PF00785; PAC.1.
 DR SMART: SM00100; CNMP.1.
 DR SMART: SM00086; PAC.1.
 DR PROSITE: PS50042; CNMP_BINDING_3; 1.
 DR PROSITE: PS50112; PAS.1.
 DR PROSITE: PS50113; PAC.1.
 KM Transport; Ion transport; Ionic channel; Voltage-gated channel;
 KM Potassium channel; Potassium; Potassium transport; Transmembrane;
 KM Phosphorylation; Glycoprotein; Multigene family; Alternative splicing.
 FT DOMAIN 1 405
 FT TRANSMEM 406 426
 FT TRANSMEM 453 473
 FT TRANSMEM 474 497
 FT TRANSMEM 498 518
 FT TRANSMEM 523 543
 FT TRANSMEM 544 549
 FT TRANSMEM 550 570
 FT TRANSMEM 614 634
 FT TRANSMEM 641 661
 FT TRANSMEM 662 1161
 FT DOMAIN 17 88
 FT DOMAIN 92 144
 FT DOMAIN 299 302
 FT NP_BIND 744 861
 FT CARBOHYD 600 600
 FT VANSPLIC 69 85
 FT CONFLICT 411 411
 FT CONFLICT 445 446
 FT CONFLICT 553 553
 FT CONFLICT 561 561
 FT SEQUENCE 1161 AA; 126961 MW; 798532B2FFBD9AEB CRC64;

Query Match 97.7%; Score 5942; DB 1; Length 1161;
 Best Local Similarity 98.0%; Pred. No. 4,1e-300;
 Matches 1138; Conservative 2; Mismatches 19; Indels 2; Gaps 1;

QY 1 MPVRGHVAPONTFLDTIIRKFEKGRFTIITANAVENCAYIYNDGFCGLGYSRAEVM 60
 DB 1 MPVRGHVAPONTFLDTIIRKFEKGRFTIITANAVENCAYIYNDGFCGLGYSRAEVM 60

QY 61 QRPCTCDLHGPRTQRRRAAIOALLGAERKVEIATYRKDSCSEFLCLVNVVPKNDG 120
 DB 61 QRPCTCDLHGPRTQRRRAAIOALLGAERKVEIATYRKDSCSEFLCLVNVVPKNDG 120

QY 121 AVIMEILNFEVMEKDMGSPAHDTNHRGPTSMILAPRAKTFRLKLALLATFARSSV 180
 DB 121 AVIMEILNFEVMEKDMGSPAHDTNHRGPTSMILAPRAKTFRLKLALLATFARSSV 180

QY 121 AVIMEILNFEVMEKDMGSPAHDTNHRGPTSMILAPRAKTFRLKLALLATFARSSV 180
 DB 121 AVIMEILNFEVMEKDMGSPAHDTNHRGPTSMILAPRAKTFRLKLALLATFARSSV 180

QY 181 RSGAGAGAGACAVVVDVLTLPAAFSSESLALDEVYANDNVHAGLGAEEERRALVPGSP 240
 DB 181 RSGAGAGAGACAVVVDVLTLPAAFSSESLALDEVYANDNVHAGLGAEEERRALVPGSP 240

QY 241 PR--SAREQLPSPRAHSLNPDASSGSCSLANTRSRESCASVRRASSADDEAMRAGVLP 298

DB 241 PPPVASGPHSPRAHSLNPDASSGSCSLANTRSRESCASVRRASSADDEAMRAGVLP 300

QY 299 PPRHASTGAMHPLRSGLLNSTSDSLVRYRTSKIPQITLNFVDLGDGPFLLASPSDBEI 358
 DB 301 PPRHASTGAMHPLRSGLLNSTSDSLVRYRTSKIPQITLNFVDLGDGPFLLASPSDBEI 360

QY 359 IAPKIKERTHNVTEKVTQVLSIGADVLEPKLQAPRIHRMTLHSPKAVWMDLILLY 418
 DB 361 IAPKIKERTHNVTEKVTQVLSIGADVLEPKLQAPRIHRMTLHSPKAVWMDLILLY 420

QY 419 IYTAFTYSAFELKTEEBGPACQPLAVVLLYDIMYVILLNFRTTYNA 478
 DB 421 IYTAFTYSAFELKTEEBGPACQPLAVVLLYDIMYVILLNFRTTYNA 480

QY 479 NEEVSHHGRIAVHFKGMFLDMVAALPFDLLIGSGSEELIGLKTARLLRLTVARK 538
 DB 481 NEEVSHHGRIAVHFKGMFLDMVAALPFDLLIGSGSEELIGLKTARLLRLTVARK 540

QY 539 LDRSEYGAVALFLMCTFALIAHWAICWAIGMEOPHMDSRIGMLHNLGDQIKDPYN 598
 DB 541 LDRSEYGAVALFLMCTFALIAHWAICWAIGMEOPHMDSRIGMLHNLGDQIKDPYN 600

QY 599 SSGLGSPSITKDYATLFTSSSLTVSGFANVSPNTSEKIFSTICVMLIGSLMTASTGN 658
 DB 601 SSGLGSPSITKDYATLFTSSSLTVSGFANVSPNTSEKIFSTICVMLIGSLMTASTGN 660

QY 659 VSAIIQRLYSGTARYHTOMLRYREFIRFHQIPNPLRQBLEEYFOHMSYTNIGDINNAVLK 718
 DB 661 VSAIIQRLYSGTARYHTOMLRYREFIRFHQIPNPLRQBLEEYFOHMSYTNIGDINNAVLK 720

QY 719 GFPECLQADICLHLNRSLLQCKPFRGATKGLRALAMKFTTAPRGGDTLVHAGDLTA 778
 DB 721 GFPECLQADICLHLNRSLLQCKPFRGATKGLRALAMKFTTAPRGGDTLVHAGDLTA 780

QY 779 LYFISRSSEILIRGVVVAIIIGKNDICEPNTLYARPKSGNDVVALYCDLHKTHRD 838
 DB 781 LYFISRSSEILIRGVVVAIIIGKNDICEPNTLYARPKSGNDVVALYCDLHKTHRD 840

QY 839 LEVLDMPEDSDHEFSSLEITFNLDNTNMIPGSPSTLEGGFSRQRRLKSFRRRTDKD 898
 DB 841 LEVLDMPEDSDHEFSSLEITFNLDNTNMIPGSPSTLEGGFSRQRRLKSFRRRTDKD 900

QY 899 TEQGEVSALGRRAGAPSSRGPRGPGWGESPPSSSEDEGGGRSSPLRLVPF 958
 DB 901 TEQGEVSALGRRAGAPSSRGPRGPGWGESPPSSSEDEGGGRSSPLRLVPF 960

QY 959 SSPRRGPRGGEPLMECEKSSDTCNPLSGAFSGVSNIFSPWGSRRGROYELPRCAP 1018
 DB 961 SSPRRGPRGGEPLMECEKSSDTCNPLSGAFSGVSNIFSPWGSRRGROYELPRCAP 1020

QY 1019 TPSSLNIPLSSPGRPRGDSRLDALQROLNRLTETRLSADMAVYLOLQOMTLVPPAY 1078
 DB 1021 TPSSLNIPLSSPGRPRGDSRLDALQROLNRLTETRLSADMAVYLOLQOMTLVPPAY 1080

QY 1079 SAVTTTGGPGPTSPPLFVSPPLTLDLSQVSOQFMAACEBLPGCAPLPDGPTRRLSL 1138
 DB 1081 SAVTTTGGPGPTSPPLFVSPPLTLDLSQVSOQFMAACEBLPGCAPLPDGPTRRLSL 1140

QY 1139 PGOLGALTSOPLHRRGSDPS 1159
 DB 1141 PGOLGALTSOPLHRRGSDPS 1161

RESULT 3
 KCH2_CANFA STANDARD; PRT; 1158 AA.
 AC Q9TSZ3; 002719; 018820;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Potassium voltage-gated channel subfamily H member 2 (Ether-a-go-go related gene potassium channel 1) (ERGI) (c-erg) (DERG) (Ether-a-go-

DE go related protein 1) (Eag related protein 1).
GN KCNH2 OR ERG OR CERG.
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX NCBI_TaxID=9615;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Heart;
RX MEDLINE=21310885; PubMed=11417212;
RA Zehetlein J., Zhang W., Koenen M., Graf M., Heinemann S.H., Katus H.A.;
RT "Molecular cloning and expression of CERG, the ether a go-go-related
gene from canine myocardium."
RL Pfugers Arch. 442:188-191(2001).
RN [2]
RP SEQUENCE OF 407-566 FROM N.A.
RC TISSUE=Heart;
RX MEDLINE=97164986; PubMed=9012748;
RA Wymore R.S., Gintant G.A., Wymore R.T., Dixon J.E., McKinnon D.,
RA Cohen I.S.;
RT "Tissue and species distribution of mRNA for the IKr-like K⁺ channel,
ERG.".
RL Circ. Res. 80:261-268(1997).
RN [3]
RP SEQUENCE OF 616-714 FROM N.A.
RC TISSUE=Heart atrium;
RX MEDLINE=99221626; PubMed=10205145;
RA Yue L., Melnyk P., Gaspo R., Wang Z., Nattel S.;
RT "Molecular mechanisms underlying ionic remodeling in a dog model of
atrial fibrillation."
RL Circ. Res. 84:776-784(1999).
CC -1- FUNCTION: Pore-forming (alpha) subunit of voltage-gated inwardly
rectifying potassium channel. Channel properties are modulated by
CAMP and subunit assembly. Mediates the rapidly activating
component of the delayed rectifying potassium current in heart
(IKr) (By similarity).
CC -1- SUBUNIT: The potassium channel is probably composed of a homo- or
heterotetrameric complex of pore-forming alpha subunits that can
associate with modulating beta subunits. Heteromultimer with
KCNH6/ERG2, KCNH7/ERG3, KCNE1 and KCNE2 (By similarity).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- TISSUE SPECIFICITY: Highly expressed in left and right atria of
the heart, in cortex and hippocampus; detected at intermediate
levels in left and right ventricle, Purkinje fibers, cerebellum,
thalamus and basal ganglia; detected at low levels in liver,
spleen and kidney.
CC -1- DOMAIN: The segment S4 is probably the voltage-sensor and is
characterized by a series of positively charged amino acids at
every third position.
CC -1- PTM: Phosphorylated on serine and threonine residues (By
similarity).
CC -1- SIMILARITY: Belongs to the potassium channel family. H (Eag)
subfamily.
CC -1- SIMILARITY: Contains 1 cyclic nucleotide-binding domain.
CC -1- SIMILARITY: Contains 1 PAS (PER-ARNT-SIM) dimerization domain.
CC -1- SIMILARITY: Contains 1 PAS-associated C-terminal (PAC) domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation-
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@sib-sib.ch).
CC -----
DR EMBL: AJ243344; CAB64868.1; -
DR EMBL: U75213; AAC48722.1; -
DR EMBL: AF017429; AAB70524.1; -
DR InterPro: IPR000595; CNMP-binding.
DR InterPro: IPR003967; Erg_channel.
DR InterPro: IPR005821; Ion_trans.
DR InterPro: IPR001622; K+channel_pore.
DR InterPro: IPR005820; M+channel_nlg.

DR InterPro: IPR001610; PAC.
DR InterPro: IPR000700; PAS-assoc_C.
DR InterPro: IPR000014; PAS_domain.
DR Pfam: PF000027; CNMP_binding; 1.
DR Pfam: PF00520; Ion_trans; 1.
DR Pfam: PF00785; PAC; 1.
DR SMART: SM00100; CNMP; 1.
DR SMART: SM00086; PAC; 1.
DR SMART: SM00091; PAS; 1.
DR PROSITE: PSS0042; CNMP_BINDING_3; 1.
DR PROSITE: PSS0112; PAS; 1.
DR PROSITE: PSS0113; PAC; 1.
DR Transport: Ion transport; Ionic channel; Voltage-gated channel;
KW Potassium channel; Potassium; Potassium transport; Transmembrane;
KW Phosphorylation; Glycoprotein; Multigene family;
FT DOMAIN 1 402
FT TRANSMEM 403 423 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 450 470 SEGMENT S1 (POTENTIAL).
FT TRANSMEM 471 494 SEGMENT S2 (POTENTIAL).
FT TRANSMEM 495 515 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 520 540 SEGMENT S3 (POTENTIAL).
FT TRANSMEM 541 546 SEGMENT S4 (POTENTIAL).
FT TRANSMEM 547 567 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 611 631 SEGMENT S5 (POTENTIAL).
FT TRANSMEM 638 658 SEGMENT H5 (PORE-FORMING) (POTENTIAL).
FT DOMAIN 659 1158 SEGMENT S6 (POTENTIAL).
FT DOMAIN 17 88 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 92 144 PAS.
FT DOMAIN 296 299 PAC.
FT NP_BIND 741 858 POLY-PRO.
FT CARBOHYD 597 597 CNMP.
FT CONFLICT 442 442 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 1158 AA; 126644 MW; 530849032B4A43D0 CRC64;

Query Match 96.68; Score 5872.5; DB 1; Length 1158;
Best Local Similarity 96.7%; Pred. No. 1.0e-296;
Matches 1121; Conservative 8; Mismatches 29; Indels 1; Gaps 1;

QY 1 MPVRGVAPONFDLTIRKFEQSQRKFIANARVNCVITCNOGFCGSRAEVM 60
1 MPRRGVAPONFDLTIRKFEQSQRKFIANARVNCVITCNOGFCGSRAEVM 60
DB 1 MPRRGVAPONFDLTIRKFEQSQRKFIANARVNCVITCNOGFCGSRAEVM 60
QY 61 QRCCTDFLHGPTORRAAQAIALGAEREVEIAFYRKDSCFCLVDPVVKNEG 120
61 QRCCTDFLHGPTORRAAQAIALGAEREVEIAFYRKDSCFCLVDPVVKNEG 120
DB 61 QRCCTDFLHGPTORRAAQAIALGAEREVEIAFYRKDSCFCLVDPVVKNEG 120
QY 121 AVIMEILNFEVMEKDWGSPAHDTNHRGPTSMLAGRAKTRFLKIPALLATARESSV 180
121 AVIMEILNFEVMEKDWGSPAHDTNHRGPTSMLAGRAKTRFLKIPALLATARESSV 180
DB 121 AVIMEILNFEVMEKDWGSPAHDTNHRGPTSMLAGRAKTRFLKIPALLATARESSV 180
QY 181 RSGGAGAGAPGAVVVDVDTTPAAPSSESLADEVTAMDNHVGAGAEERRALVGGSP 240
181 RSGGAGAGAPGAVVVDVDTTPAAPSSESLADEVTAMDNHVGAGAEERRALVGGSP 240
DB 181 RSGGAGAGAPGAVVVDVDTTPAAPSSESLADEVTAMDNHVGAGAEERRALVGGSP 240
QY 241 PRAPGOLPPRRHSLNLPDASGSSCLARTRRESCASVRRASADIDEAMRAGVLP 300
241 PRAPGOLPPRRHSLNLPDASGSSCLARTRRESCASVRRASADIDEAMRAGVLP 300
DB 241 PRAPGOLPPRRHSLNLPDASGSSCLARTRRESCASVRRASADIDEAMRAGVLP 300
QY 241 PAGAPRPLSPRRHSLNLPDASGSSCLARTRRESCASVRRASADIDEAMRAGVLP 299
241 PAGAPRPLSPRRHSLNLPDASGSSCLARTRRESCASVRRASADIDEAMRAGVLP 299
DB 241 PAGAPRPLSPRRHSLNLPDASGSSCLARTRRESCASVRRASADIDEAMRAGVLP 299
QY 301 RHASTGAMHPLRGLNLSTSDSLVRYRTISKIPQITLNFVDLKGDFLSPSDREIIA 360
301 RHASTGAMHPLRGLNLSTSDSLVRYRTISKIPQITLNFVDLKGDFLSPSDREIIA 360
DB 301 RHASTGAMHPLRGLNLSTSDSLVRYRTISKIPQITLNFVDLKGDFLSPSDREIIA 360
QY 300 RHASTGAMHPLRGLNLSTSDSLVRYRTISKIPQITLNFVDLKGDFLSPSDREIIA 359
300 RHASTGAMHPLRGLNLSTSDSLVRYRTISKIPQITLNFVDLKGDFLSPSDREIIA 359
DB 300 RHASTGAMHPLRGLNLSTSDSLVRYRTISKIPQITLNFVDLKGDFLSPSDREIIA 359
QY 361 PKIKERHNTKVTQVLSGADVLPYKIQAPRIHWTILHSPFAVMDMLLLVIY 420
361 PKIKERHNTKVTQVLSGADVLPYKIQAPRIHWTILHSPFAVMDMLLLVIY 420
DB 361 PKIKERHNTKVTQVLSGADVLPYKIQAPRIHWTILHSPFAVMDMLLLVIY 420
QY 360 PKIKERHNTKVTQVLSGADVLPYKIQAPRIHWTILHSPFAVMDMLLLVIY 419
360 PKIKERHNTKVTQVLSGADVLPYKIQAPRIHWTILHSPFAVMDMLLLVIY 419
DB 360 PKIKERHNTKVTQVLSGADVLPYKIQAPRIHWTILHSPFAVMDMLLLVIY 419
QY 421 TAVFTYSAFLKETEEGPPATECGYACOPAVVDLIDVIMFDLINFRTTYVANE 480
421 TAVFTYSAFLKETEEGPPATECGYACOPAVVDLIDVIMFDLINFRTTYVANE 480
DB 421 TAVFTYSAFLKETEEGPPATECGYACOPAVVDLIDVIMFDLINFRTTYVANE 480
QY 420 TAVFTYSAFLKETEEGPPADCGAACOPAVVDLIDVIMFDLINFRTTYVANE 479
420 TAVFTYSAFLKETEEGPPADCGAACOPAVVDLIDVIMFDLINFRTTYVANE 479
DB 420 TAVFTYSAFLKETEEGPPADCGAACOPAVVDLIDVIMFDLINFRTTYVANE 479
QY 481 EVVSHPGRIAVHFKGFLIDVYAAIPFDLLIFGSGSEELIGLTKTARLLRVARKLD 540
481 EVVSHPGRIAVHFKGFLIDVYAAIPFDLLIFGSGSEELIGLTKTARLLRVARKLD 540
DB 481 EVVSHPGRIAVHFKGFLIDVYAAIPFDLLIFGSGSEELIGLTKTARLLRVARKLD 540
QY 480 EVVSHPGRIAVHFKGFLIDVYAAIPFDLLIFGSGSEELIGLTKTARLLRVARKLD 539
480 EVVSHPGRIAVHFKGFLIDVYAAIPFDLLIFGSGSEELIGLTKTARLLRVARKLD 539
DB 480 EVVSHPGRIAVHFKGFLIDVYAAIPFDLLIFGSGSEELIGLTKTARLLRVARKLD 539

QY 541 RYSEYGAAYFLMCTFALLAHMLACIWAYAIGNMEOPHMDSRIGLNLHNGDQIGKPNSS 600
 DB 540 RYSEYGAAYFLMCTFALLAHMLACIWAYAIGNMEOPHMDSRIGLNLHNGDQIGKPNSS 599
 QY 601 GLGSPSTKDKYVYALYFFSSLSVGRGNVSPNTNSEKIFSCWLGSLMYASTFQVNS 660
 DB 600 GLGSPSTKDKYVYALYFFSSLSVGRGNVSPNTNSEKIFSCWLGSLMYASTFQVNS 659
 QY 661 AIIORLXSGTARYTOMLRVREFIRFHOIPNPLRQRLREYEQHMASTYNGIDMNAVLKGF 720
 DB 660 AIIORLXSGTARYTOMLRVREFIRFHOIPNPLRQRLREYEQHMASTYNGIDMNAVLKGF 719
 QY 721 PECLQADICHLNLSLQHCCKPFRGATKCLRALAMKFKTHAPDGLVHAGDGLTALY 780
 DB 720 PECLQADICHLNLSLQHCCKPFRGATKCLRALAMKFKTHAPDGLVHAGDGLTALY 779
 QY 781 FISGSIETLRGDVVAALGKNDIFGEPLNIYARPGKNGVRYALTYCDLKHRRDLE 840
 DB 780 FISGSIETLRGDVVAALGKNDIFGEPLNIYARPGKNGVRYALTYCDLKHRRDLE 839
 QY 841 VLDMYPERSDHFMSSLETTFNLRDTNMIPGSPGSTTELEGFSRORRKLSPRRRTDKDTE 900
 DB 840 VLDMYPERSDHFMSSLETTFNLRDTNMIPGSPGSTTELEGFSRORRKLSPRRRTDKDTE 899
 QY 901 QPGEVSALGPGRAGAPSSSRGRPGCPWGESPPSSSGPSSSESDGCPGRSSSPLRLVPRSS 960
 DB 900 QPGEVSALGPGRAGAPSSSRGRPGCPWGESPPSSSGPSSSESDGCPGRSSSPLRLVPRSS 959
 QY 961 PRPGEPPGGPRLMEDCKSSDTCNPISGAFSGVSNIFSPMGDSRGROYELPRCPATP 1020
 DB 960 PRPGEPPGGPRLMEDCKSSDTCNPISGAFSGVSNIFSPMGDSRGROYELPRCPATP 1019
 QY 1021 SLNIPILSSPPRRRGDVEESRLDALQRLNLETRLSDMATVYQLORQTVLPAYSA 1080
 DB 1020 SLNIPILSSPPRRRGDVEESRLDALQRLNLETRLSDMATVYQLORQTVLPAYSA 1079
 QY 1081 VTPPGPPTSTPLPVSPPTLTLDLSISOVSQFMACEELPPGAPBELPQEGPTRLSLPG 1140
 DB 1080 VTPPGPPTSTPLPVSPPTLTLDLSISOVSQFMACEELPPGAPBELPQEGPTRLSLPG 1139
 QY 1141 QLGLATSQPLHRHSGDPGS 1159
 DB 1140 QLGLATSQPLHRHSGDPGS 1158

RESULT 4

KCH2_MOUSE STANDARD; PRT: 1162 AA.
 AC 035219; 035220; 035221; 035989;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Potassium voltage-gated channel subfamily H member 2 (Ether-a-go-go related gene potassium channel 1) (ERGL1) (MERG1) (Ether-a-go-go related protein 1) (Eag related protein 1).
 GN KCHN2 OR ERG OR MERG1.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NC NCB1_TaxID=10090;
 OX [1]
 RN
 RP SEQUENCE FROM N.A. (ISOFORMS 1; 2 AND 3), AND VARIANTS ARG-186; THR-455; TYR-752 AND ASN-1006.
 RC STRAIN=129/SV, and BALB/c; TISSUE=Heart;
 RX MEDLINE=98012815; PubMed=9351462;
 RA London B., Trudeau M.C., Newton K.P., Beyer A.K., Copeland N.G., Gilbert D.J., Jenkins N.A., Sailer C.A., Robertson G.A.;
 RT "Two isoforms of the mouse ether-a-go-go-related gene coassemble to form channels with properties similar to the rapidly activating RT component of the cardiac delayed rectifier K⁺ current.";
 RL Circ. Res. 81:870-878(1997).
 RN [2]

RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 3).
 RC TISSUE=atrial tumor;
 RX MEDLINE=98012799; PubMed=9351446;
 RA Lees-Miller J.P., Rondo C., Wang L., Duff H.J.;
 RT "Electrophysiological characterization of an alternatively processed ERG K⁺ channel in mouse and human hearts.";
 RL Circ. Res. 81:719-726(1997).
 RN [3]
 RP SEQUENCE FROM N.A. (ISOFORM 1), AND VARIANT ARG-186.
 RC TISSUE=Colon smooth muscle;
 RA Shoen F., Malykhina A., Akbarali H.I.;
 RT "Smooth muscle KCHN2 erg potassium channel.";
 RL Submitted (JUL-2002) to the EMBL/Genbank/DBJ databases.
 CC -1- FUNCTION: Pore-forming (alpha) subunit of voltage-gated inwardly rectifying potassium channel. Channel properties are modulated by CAMP and subunit assembly. Mediates the rapidly activating component of the delayed rectifying potassium current in heart (IKr) (By similarity).
 CC -1- SUBUNIT: The potassium channel is probably composed of a homo- or heterotetrameric complex of pore-forming alpha subunits that can associate with modulating beta subunits. Heteromultimer with KCNH6/ERG2, KCNH7/ERG3, KCNE1 and KCNE2 (By similarity).
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- ALTERNATIVE PRODUCTS: Integral membrane protein.
 CC Event=Alternative splicing; Named isoforms=3;
 CC Comment=Experimental confirmation may be lacking for some isoforms;
 CC Name=1; Synonyms=1A, A;
 CC IsoId=035219-1; Sequence=Displayed;
 CC Name=2; Synonyms=1A';
 CC IsoId=035219-2; Sequence=VSP_000969;
 CC Name=3; Synonyms=1B, B;
 CC IsoId=035219-3; Sequence=VSP_000970;
 CC -1- TISSUE SPECIFICITY: Isoform 1 is expressed in heart, brain and testis and at low levels in lung. Isoform 3 is expressed predominantly in heart. The expression of isoform 2 is low in all tissues tested.
 CC -1- DOMAIN: The segment S4 is probably the voltage-sensor and is characterized by a series of positively charged amino acids at every third position.
 CC -1- PTM: Phosphorylated on serine and threonine residues (By similarity).
 CC -1- SIMILARITY: Belongs to the potassium channel family. H (Eag) subfamily.
 CC -1- SIMILARITY: Contains 1 cyclic nucleotide-binding domain.
 CC -1- SIMILARITY: Contains 1 PAS (PER-ARNT-SIM) dimerization domain.
 CC -1- SIMILARITY: Contains 1 PAS-associated C-terminal (PAC) domain.
 CC -1- CAUTION: Ref.3 sequence was originally reported as deriving from rabbit.
 CC -1- CAUTION: Ref.3 sequence differs from that shown due to a frameshift in position 1057.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AF012868; AAC3418.1; -;
 DR EMBL; AF012869; AAC3419.1; -;
 DR EMBL; AF012871; AAC3420.1; -;
 DR EMBL; AF012870; AAC3420.1; JOINED.
 DR EMBL; AF012871; AAC3421.1; -;
 DR EMBL; AF012871; AAC3422.1; -;
 DR EMBL; AF012870; AAC3422.1; JOINED.
 DR EMBL; AF034762; AAB87571.1; -;
 DR EMBL; AF439342; AAL35337.2; ALT_FRAME.
 DR MGD; MGI:1341722; Kcnh2.
 DR InterPro; IPR000595; CNMP_binding.
 DR InterPro; IPR003967; Erg_channel.
 DR InterPro; IPR005821; Ion_trans.

DR InterPro: IPR001622; K+channel_pore.
 DR InterPro: IPR005820; M+channel_nlg.
 DR InterPro: IPR001610; PAC.
 DR InterPro: IPR000700; PAS-associ_C.
 DR InterPro: IPR000014; PAS_domain.
 DR Pfam: PF00027; CNMP_binding_1.
 DR Pfam: PF00520; ion_trans_1.
 DR Pfam: PF00785; PAC_1.
 DR SMART: SM00100; CNMP_1.
 DR SMART: SM00086; PAC_1.
 DR PROSITE: PS50042; CNMP_BINDING_3_1.
 DR PROSITE: PS50112; PAC_1.
 DR PROSITE: PS50113; PAC_1.
 KW Transport; Ion transport; Potassium channel; Voltage-gated membrane;
 KW Potassium channel; Potassium; Potassium transport; Transmembrane;
 KW Phosphorylation; Glycoprotein; Multigene family; Polymorphism;
 KW Alternative splicing.
 FT DOMAIN 1 405
 FT TRANSMEM 406 426 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 453 473 SEGMENT S1 (POTENTIAL).
 FT DOMAIN 474 497 SEGMENT S2 (POTENTIAL).
 FT TRANSMEM 498 518 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 523 543 SEGMENT S3 (POTENTIAL).
 FT TRANSMEM 544 549 SEGMENT S4 (POTENTIAL).
 FT TRANSMEM 550 570 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 614 634 SEGMENT S5 (POTENTIAL).
 FT TRANSMEM 641 661 SEGMENT S6 (POTENTIAL).
 FT TRANSMEM 662 1162 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 17 88
 FT DOMAIN 92 144 PAC.
 FT DOMAIN 299 302 POLY-PRO.
 FT NP_BIND 744 861
 FT CARBOHYD 600 600
 FT VARSPIC 1 59
 FT VARSPIC 1 378
 FT VARSPIC 1 378
 FT MPARGHVAPONTFLDTIIRKFEQSRKFTIINARVENCAY
 IYCNDFGCELCYGRAEMORPCYDGLHGPRTORRAAOI
 AOALLGAERKVEIAFYRKDSCFLCLVDVAVYKNEGAYI
 MFLNFEVMEKDMVGPAPDTHNRGPTSMLAGRAKTR
 LKIPALLALTARESSVYTGSMHSGAGAVYVDYDLPAAP
 SSESALDEVAMDNHAGLPAERLALVPGSASVVASI
 RGPSPRQSLNDAGSSCSLARTSRSCASVRSASSA
 DLEAMRAGALPPPRHASTGAMHPLSGILNSTDYLR
 YRTISKIPOLTNVDKDPFLASPTSDREIIPKIKERT
 HNYEKYTO -> MAIPTEKSRGTALQPRAGCRVARRAV
 ISSLVAGE (in isoform 3).
 /FtId=VSP_000970.
 FT VARIANT 186 186
 FT VARIANT 455 455 A -> R (IN STRAIN BALB/C).
 FT VARIANT 752 752 C -> Y (IN STRAIN BALB/C).
 FT VARIANT 1006 1006 D -> N (IN STRAIN BALB/C).
 FT CONFLICT 111 111 D -> G (IN REF. 3).
 FT CONFLICT 126 126 I -> V (IN REF. 3).
 FT CONFLICT 198 198 V -> A (IN REF. 3).
 FT CONFLICT 214 214 E -> G (IN REF. 3).
 FT CONFLICT 537 537 V -> A (IN REF. 3).
 FT CONFLICT 908 908 S -> P (IN REF. 3).
 FT CONFLICT 929 929 G -> R (IN REF. 3).
 SQ SEQUENCE 1162 AA; 126885 MM; A9455F7F10B61E46 CRC64;
 Query Match 95.4%; Score 5797.5; DB 1; Length 1162;
 Best Local Similarity 96.0%; Pred. No. 1.2e-292;
 Matches 116; Conservative 10; Mismatches 32; Indels 5; Gaps 3;

DB 121 AVIMFLNFEVMEKDMVGPAPDTHNRGPTSMLAGRAKTRFLKLPALLALTARESSV 180
 QY 181 RSGAGAGAGAAVVDVLTLPAPSSSESLALDEVYAMDHNHAGLPAERLALVPG-- 238
 DB 181 RTGSMHSGAGAAVVDVLTLPAPSSSESLALDEVYAMDHNHAGLPAERLALVPGSA 240
 QY 239 SPFRSARAGLPSPRAHSLNDPAGSSCSLARTSRSCASVRRASSADDI EAMAGVLP 298
 DB 241 SPVASIRGPHSPSPAGSLNDPAGSSCSLARTSRSCASVRRASSADDI EAMAGVLP 300
 QY 299 PPRHASTGAMHPLRSGILNSTDSDLYRYRTISKIPQITLNFVDLKDPLASPTSOR 358
 DB 301 PPRHASTGAMHPLRSGILNSTDSDLYRYRTISKIPQITLNFVDLKDPLASPTSOR 360
 QY 359 IAPKIKERTHNYEKYTOVLSLGDVLPYKLAQPRHRTILHYSPPKAVMDLILLY 418
 DB 361 IAPKIKERTHNYEKYTOVLSLGDVLPYKLAQPRHRTILHYSPPKAVMDLILLY 420
 QY 419 IYTAVFPTPSAFLKTEEGPATEGCGYACQPLAVVDLIVDIMEIVDILINPRTYVNA 478
 DB 421 IYTAVFPTPSAFLKTEEGSQAPDCGYACQPLAVVDLIVDIMEIVDILINPRTYVNA 480
 QY 479 NEEVSHRGRIVAHYFKGWEFLIDVNAIIPFDLIFGSGSEELIGLKTARLLRYVARK 538
 DB 481 NEEVSHRGRIVAHYFKGWEFLIDVNAIIPFDLIFGSGSEELIGLKTARLLRYVARK 540
 QY 539 LDYSEYGAAVFLMCTFALLIAHWLACIWAIGNMEQPHMDSRIGLHNLGDOIGKPN 598
 DB 541 LDYSEYGAAVFLMCTFALLIAHWLACIWAIGNMEQPHMDSRIGLHNLGDOIGKPN 600
 QY 599 SSGIGGSPSIDKYTALYFTFESSLTSGFGVSNPTSEKIFSCVWLIGSIMASTFEN 658
 DB 601 SSGIGGSPSIDKYTALYFTFESSLTSGFGVSNPTSEKIFSCVWLIGSIMASTFEN 660
 QY 659 VSAIIQGLYSGTARYHTOMLRVREFIRFHQIIPNLRORLEYPOHASTYNGIDMNVLK 718
 DB 661 VSAIIQGLYSGTARYHTOMLRVREFIRFHQIIPNLRORLEYPOHASTYNGIDMNVLK 720
 QY 719 GFPECLDADICLHNSRLDHCXPFPGATYGCCLALMKRKTTHAPPGDTLVHAGDLLTA 778
 DB 721 GFPECLDADICLHNSRLDHCXPFPGATYGCCLALMKRKTTHAPPGDTLVHAGDLLTA 780
 QY 779 LYFISRSISIELRQDVVVALIGKNDIRGEPLNLYARPGKNSGDVRLTYCDLKHTRDD 838
 DB 781 LYFISRSISIELRQDVVVALIGKNDIRGEPLNLYARPGKNSGDVRLTYCDLKHTRDD 840
 QY 839 LEVLDMPPEFSDHFWSSLETFNLDTNMIIPGSPGSTELGSGFSRQRRKLSFFRRDKD 898
 DB 841 LEVLDMPPEFSDHFWSSLETFNLDTNMIIPGSPGSTELGSGFSRQRRKLSFFRRDKD 900
 QY 899 TEOPGEVSAI--GPRGAGBPSSRGPRGPGWGSPPSSGESSDEDEGPRSSPLRLV 956
 DB 901 TEOPGEVSAI--GPRGAGBPSSRGPRGPGWGSPPSSGESSDEDEGPRSSPLRLV 960
 QY 957 PFSSPRPGPGPGGPELMEDECKSSDFCNPLSGAFSGVSNFSGWGSRGROYOELRCP 1016
 DB 961 PFSSPRPGPGPGGPELMEDECKSSDFCNPLSGAFSGVSNFSGWGSRGROYOELRCP 1019
 QY 1017 APTPSLNLPLSPGRRRGDVEDSRDLALROLRLTETRLSADANATYLOLQOROMTLVPP 1076
 DB 1020 APTPSLNLPLSPGRRRGDVEDSRDLALROLRLTETRLSADANATYLOLQOROMTLVPP 1079
 QY 1077 AYSAVTTTPGPGPSTSLPLVSPPLTLTDSLOVSQPMACBELPGAPLPOEGPTRL 1136
 DB 1080 AYSAVTTTPGPGPSTSLPLVSPPLTLTDSLOVSQPMACBELPGAPLPOEGPTRL 1139
 QY 1137 SLPGOLGALTSLPLHHRGSDPGS 1159
 DB 1140 SLPGOLGALTSLPLHHRGSDPGS 1162
 RESULT 5

KCH2_RAT STANDARD; PRT; 1163 AA.
 ID KCH2_RAT 008962; 008720;
 AC 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Potassium voltage-gated channel subfamily H member 2 (Ether-a-go-go related gene potassium channel 1) (ERG1) (r-ERG) (RERG) (Ether-a-go-go related protein 1) (Eag related protein 1).
 GN KCHN2 OR ERG.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain cortex;
 RX MEDLINE=98329322; PubMed=9664620;
 RA Bauer C.K., Engeland B., Wulfsen I., Ludwig J., Pongs O., Schwarz J.R.;
 RT "RERG is a molecular correlate of the inward-rectifying K current in clonal net pituitary cells."
 RL Recept. Channels 6:19-29(1998).
 RN [2]
 RP SEQUENCE OF 409-568 FROM N.A.
 RX MEDLINE=97164986; PubMed=9012748;
 RA Wymore R.S., Gintant G.A., Wymore R.T., Dixon J.E., McKinnon D., Cohen I.S.;
 RT "Tissue and species distribution of mRNA for the IKr-like K+ channel, erg."
 RL Circ. Res. 80:261-268(1997).
 RN [3]
 RP TISSUE SPECIFICITY.
 RX MEDLINE=20183472; PubMed=10718922;
 RA Wulfsen I., Hauber H.P., Schiemann D., Bauer C.K., Schwarz J.R.;
 RT "Expression of mRNA for voltage-dependent and inward-rectifying K channels in GH3/86 cells and rat pituitary."
 RL J. Neuroendocrinol. 12:263-272(2000).
 RN [4]
 RP INTERACTION WITH KCHN6 AND KCHN7, AND MUTAGENESIS OF GLY-630.
 RX MEDLINE=21079731; PubMed=11212207;
 RA Wimmers S., Wulfsen I., Bauer C.K., Schwarz J.R.;
 RT "Erg1, erg2 and erg3 K channel subunits are able to form heteromultimers."
 RL Pfluegers Arch. 441:450-455(2001).
 CC -1- FUNCTION: Pore-forming (alpha) subunit of voltage-gated inwardly rectifying potassium channel. Channel properties are modulated by CAMP and subunit assembly. Mediates the rapidly activating component of the delayed rectifying potassium current in heart (IKr) (By similarity).
 CC -1- SUBUNIT: The potassium channel is probably composed of a homo- or heterotetrameric complex of pore-forming alpha subunits that can associate with modulating beta subunits. Heteromultimer with KCHN6/ERG2 and KCHN7/ERG3. Heteromultimer with KCHN1 and KCHN2 (By similarity).
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- TISSUE SPECIFICITY: Highly expressed in brain and testis, slightly less so in heart, adrenal, retina and thymus. Detected at lower levels in lung, soleus, tibialis, and at very low levels in cornea and lens. A shorter transcript is detected in skeletal muscle. Found in pituitary.
 CC -1- DOMAIN: The segment S4 is probably the voltage-sensor and is characterized by a series of positively charged amino acids at every third position.
 CC -1- PTM: Phosphorylated on serine and threonine residues (By similarity).
 CC -1- SIMILARITY: Belongs to the potassium channel family. H (Eag) subfamily.
 CC -1- SIMILARITY: Contains 1 cyclic nucleotide-binding domain.
 CC -1- SIMILARITY: Contains 1 PAS (PER-ARNT-SIM) dimerization domain.
 CC -1- SIMILARITY: Contains 1 PAS-associated C-terminal (PAC) domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL; Z96106; CAB09536.1; -;
 DR EMBL; U75210; AAC53160.1; -;
 DR InterPro: IPR000595; CNMP_binding.
 DR InterPro: IPR003967; Erg_channel.
 DR InterPro: IPR005821; Ion_trans.
 DR InterPro: IPR001622; K_channel_pore.
 DR InterPro: IPR003820; M-channel_n19.
 DR InterPro: IPR001610; PAC.
 DR InterPro: IPR000700; PAS_assoc_C.
 DR InterPro: IPR000014; PAS_domain.
 DR Pfam; PF00027; CNMP_binding; 1.
 DR Pfam; PF00520; Ion_trans; 1.
 DR Pfam; PF00785; PAC; 1.
 DR SMART; SM00100; CNMP; 1.
 DR SMART; SM00086; PAC; 1.
 DR SMART; SM00091; PAS; 1.
 DR PROSITE; PS50042; CNMP_BINDING_3; 1.
 DR PROSITE; PS50112; PAS; 1.
 DR PROSITE; PS50113; PAC; 1.
 DR Transport; Ion transport; Ionic channel; Voltage-gated channel;
 KW Potassium channel; Potassium; Potassium transport; Transmembrane;
 KW Phosphorylation; Glycoprotein; Multigene family.
 FT DOMAIN 1 405
 FT TRANSMEM 406 426
 FT TRANSMEM 453 473
 FT DOMAIN 474 497
 FT TRANSMEM 498 518
 FT TRANSMEM 523 543
 FT DOMAIN 544 549
 FT TRANSMEM 550 570
 FT DOMAIN 614 634
 FT TRANSMEM 641 661
 FT DOMAIN 662 1163
 FT DOMAIN 17 88
 FT DOMAIN 92 144
 FT NP_BIND 744 861
 FT CARBOHYD 600 600
 FT MUTAGEN 630 630
 FT CONFLICT 411 411
 FT SEQUENCE 1163 AA; 126951 MW; F0D75B0B532D9EA2 CRC64;
 SQ
 Query Match 95.2%; Score 5790; DB 1; Length 1163;
 Best Local Similarity 95.6%; Pred. No. 3e-292;
 Matches 1112; Conservative 12; Mismatches 35; Indels 4; Gaps 2;
 QY 1 MPVRGHPAONTFEDTIRREKESRKEFIIANAVENCAVIYCNDSGCEICGYSRAVYM 60
 DB 1 MPVRGHPAONTFEDTIRREKESRKEFIIANAVENCAVIYCNDSGCEICGYSRAVYM 60
 QY 61 QRPCTDPLHBPRTQRRRAAQAIALGAERKEVEIAFYRRKDSGCEICLVVVPYKNEDG 120
 DB 61 QRPCTDPLHBPRTQRRRAAQAIALGAERKEVEIAFYRRKDSGCEICLVVVPYKNEDG 120
 QY 121 AVIMEILNFEVYMEKDWGSPAHDTNHRGPPSWLAPERAKTFRLKLPALLATAREBSY 180
 DB 121 AVIMEILNFEVYMEKDWGSPAHDTNHRGPPSWLAPERAKTFRLKLPALLATAREBSY 180
 QY 121 AVIMEILNFEVYMEKDWGSPAHDTNHRGPPSWLAPERAKTFRLKLPALLATAREBSY 180
 DB 121 AVIMEILNFEVYMEKDWGSPAHDTNHRGPPSWLAPERAKTFRLKLPALLATAREBSY 180
 QY 181 RSGAGAGAGAVVVDVLTLPAPSSSLALDEVTADNHNVAGIGPAEERRALVGP--G 238
 DB 181 RSGAGAGAGAVVVDVLTLPAPSSSLALDEVTADNHNVAGIGPAEERRALVGP--G 238
 QY 181 RTGSGSPGAPGAVVVDVLTLPAPSSSLALDEVTADNHNVAGIGPAEERRALVGP--G 240
 DB 181 RTGSGSPGAPGAVVVDVLTLPAPSSSLALDEVTADNHNVAGIGPAEERRALVGP--G 240
 QY 239 SPPRSAPQLPSPRAHSLNPDASGSSCLATRSRSCASVYRRASADDIEMRAGVLP 298
 DB 239 SPPRSAPQLPSPRAHSLNPDASGSSCLATRSRSCASVYRRASADDIEMRAGVLP 298
 QY 241 SPVASIRPHRSPRAQSLNPDASGSSCLATRSRSCASVYRRASADDIEMRAGVLP 300
 DB 241 SPVASIRPHRSPRAQSLNPDASGSSCLATRSRSCASVYRRASADDIEMRAGVLP 300

QY 299 PPRHASTGAMHPLRSGLNSTSDSDLVKRTKISKIPQITLNFVLDKGPFLASPTSDREI 358
 |||||
 Db 301 PPRHASTGAMHPLRSGLNSTSDSDLVKRTKISKIPQITLNFVLDKGPFLASPTSDREI 360
 |||||
 QY 359 IAKIKERHNHVEKTYOVLISGADVLPEKLOAPRIHRTILHYSFPAWMLILLY 418
 |||||
 Db 361 IAKIKERHNHVEKTYOVLISGADVLPEKLOAPRIHRTILHYSFPAWMLILLY 420
 |||||
 QY 419 IYNAVTPYSAFLKETEGEPATEGACOPFLAVVDLVDIMFVLDILINFRITYVNA 478
 |||||
 Db 421 IYNAVTPYSAFLKETEGESQAPDCGACOPFLAVVDLVDIMFVLDILINFRITYVNA 480
 |||||
 QY 479 NEEVSHPGRIAVHYFKGWFLIDMVAIPEDLLIFSGSEBELGLKTARLRLVYARK 538
 |||||
 Db 481 NEEVSHPGRIAVHYFKGWFLIDMVAIPEDLLIFSGSEBELGLKTARLRLVYARK 540
 |||||
 QY 539 LDRYSRYGAAYFLMCTPALLAHMLACTIYAGNNEQPMDSRIGMLHGLQIGKPYN 598
 |||||
 Db 541 LDRYSRYGAAYFLMCTPALLAHMLACTIYAGNNEQPMDSRIGMLHGLQIGKPYN 600
 |||||
 QY 599 SSGIGGPRISKDKYVTAIFYFTSSLTSGFGNVPNTNSEKIFISICVMLIGSLMYASIFGN 658
 |||||
 Db 601 SSGIGGPRISKDKYVTAIFYFTSSLTSGFGNVPNTNSEKIFISICVMLIGSLMYASIFGN 660
 |||||
 QY 659 VSAIIORLYSGTARYHTQMLRVREFTIRFQIIPNPLRQRLSEYPOHAMSTYNGIDMNAVLYK 718
 |||||
 Db 661 VSAIIORLYSGTARYHTQMLRVREFTIRFQIIPNPLRQRLSEYPOHAMSTYNGIDMNAVLYK 720
 |||||
 QY 719 GPEPCCQADICHLNLSLLOHCKPFRGATKGCCLALAMFKTTHARPDTLVHAGDLLTA 778
 |||||
 Db 721 GPEPCCQADICHLNLSLLOHCKPFRGATKGCCLALAMFKTTHARPDTLVHAGDLLTA 780
 |||||
 QY 779 LVFISGSIIEILRGDVVAIILGKNDIFGEPLNIYARPKSGNGVRLATYCDLKHIRDDL 838
 |||||
 Db 781 LVFISGSIIEILRGDVVAIILGKNDIFGEPLNIYARPKSGNGVRLATYCDLKHIRDDL 840
 |||||
 QY 839 LEVLDMYPERSDHFWSSLEITFENLRDTNMLPGSPGSTLEGGFSROKRLSPRRRTDKD 898
 |||||
 Db 841 LEVLDMYPERSDHFWSSLEITFENLRDTNMLPGSPGSTLEGGFSROKRLSPRRRTDKD 900
 |||||
 QY 899 TRQPGVSAI--GPGNAGAGPSSRGPRGPGWGPSSGSSPSSSEDEGGRSSPLRLV 956
 |||||
 Db 901 TRQPGVSAI--GPGNAGAGPSSRGPRGPGWGPSSGSSPSSSEDEGGRSSPLRLV 960
 |||||
 QY 957 PFSSPPPGPPGPGPELMEDCEKSSDTCNPLSGAFSGVSISEFWGDSRGROQELPRCP 1016
 |||||
 Db 961 PFSSPPPGPPGPGPELMEDCEKSSDTCNPLSGAFSGVSISEFWGDSRGROQELPRCP 1020
 |||||
 QY 1017 APPSSLNIPLSSPGRRPRGVDVSRDLAQRLNRLRLETRLSADMATVQLQRLQOMTLVPP 1076
 |||||
 Db 1021 APPSSLNIPLSSPGRRPRGVDVSRDLAQRLNRLRLETRLSADMATVQLQRLQOMTLVPP 1080
 |||||
 QY 1077 AYSAVTTPRGPRSTSPPLPVSPPLPVLTLTDSLVSQVQFMACELPFGCAPLPDEGPTTRL 1136
 |||||
 Db 1081 AYSAVTTPRGPRSTSPPLPVSPPLPVLTLTDSLVSQVQFMACELPFGCAPLPDEGPTTRL 1140
 |||||
 QY 1137 SLPGQALTSQPLHHRGSDPGS 1159
 |||||
 Db 1141 SLPGQALTSQPLHHRGSDPGS 1163
 |||||

RESULT 6
 KCH7_RAT
 ID 054852; STANDARD; PRT; 1195 AA.
 AC 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Potassium voltage-gated channel subfamily H member 7 (Ether-a-go-go related gene potassium channel 3) (Ether-a-go-go related protein 3)
 DE (Eag related protein 3).
 DE KCH7 OR ERG3.
 GN Rattus norvegicus (Rat).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=98054206; PubMed=9390998;
 RA Shi W., Wymore R.S., Wang H.S., Pan Z., Cohen I.S., McKinnon D.,
 RA Dixon J.E.;
 RT "Identification of two nervous system-specific members of the erg
 RT potassium channel gene family.";
 RL J. Neurosci. 17:9423-9432(1997).
 RN [2]
 RP TISSUE SPECIFICITY.
 RX MEDLINE=20183472; PubMed=10718922;
 RA Wolfen I., Hauber H.P., Schiemann D., Bauer C.K., Schwarz J.R.;
 RT "Expression of mRNA for voltage-dependent and inward-rectifying K
 RT channels in GH3/B6 cells and rat pituitary.";
 RL J. Neuroendocrinol. 12:263-272(2000).
 RN [3]
 RP INTERACTION WITH KCNH2 AND KCNH6, AND MUTAGENESIS OF GLY-480.
 RX MEDLINE=21079731; PubMed=11212207;
 RA Wimmers S., Wolfen I., Bauer C.K., Schwarz J.R.;
 RT "Erg1, erg2 and erg3 K channel subunits are able to form
 RT heteromultimers.";
 RL Pflugers Arch. 441:450-455(2001).
 CC -1- FUNCTION: Pore-forming (alpha) subunit of voltage-gated potassium
 CC channel. Channel properties may be modulated by
 CC CAMP and subunit assembly.
 CC -1- SUBUNIT: The potassium channel is probably composed of a homo- or
 CC heterotetrameric complex of pore-forming alpha subunits that can
 CC associate with modulating beta subunits. Heteromultimer with
 CC KCNH2/ERG1 and KCNH6/ERG2.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- TISSUE SPECIFICITY: Detected in total brain, in superior cervical,
 CC mesenteric and celiac ganglia, and at very low levels in retina.
 CC Found in pituitary.
 CC -1- DOMAIN: The segment S4 is probably the voltage-sensor and is
 CC characterized by a series of positively charged amino acids at
 CC every third position.
 CC -1- SIMILARITY: Belongs to the potassium channel family. H (Eag)
 CC subfamily.
 CC -1- SIMILARITY: Contains 1 cyclic nucleotide-binding domain.
 CC -1- SIMILARITY: Contains 1 PAS (PER-ARNT-SIM) dimerization domain.
 CC -1- SIMILARITY: Contains 1 PAS-associated C-terminal (PAC) domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: AF016191; AAB94741.1; -
 CC InterPro: IPR000595; CNMP_binding.
 CC InterPro: IPR003967; Erg_channel.
 CC InterPro: IPR005821; Ion_trans.
 CC InterPro: IPR001622; K+channel_pore.
 CC InterPro: IPR005820; M+channel_n19.
 CC InterPro: IPR001610; PAC.
 CC InterPro: IPR000700; PAS-assoc-C.
 CC InterPro: IPR000014; PAS_domain.
 CC Pfam: PF000027; CNMP_binding.1.
 CC Pfam: PF00520; ion_trans.1.
 CC Pfam: PF00785; PAC.1.
 CC SMART: SM00100; CNMP.1.
 CC SMART: SM00086; PAC.1.
 CC SMART: SM00091; PAS.1.
 CC PROSITE: PS50042; CNMP_BINDING.3; 1.
 CC PROSITE: PS50113; PAC_FALSE_NEG.
 CC PROSITE: PS50112; PAS_FALSE_NEG.
 CC Transport; Ion transport; Ionic channel; Voltage-gated channel;
 KW Potassium channel; Potassium; Potassium transport; Transmembrane;

KW Glycoprotein: Multigene family.
 FT DOMAIN 1 412 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 413 433 SEGMENT S1 (POTENTIAL).
 FT TRANSMEM 434 450 SEGMENT S2 (POTENTIAL).
 FT DOMAIN 451 470 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 471 494 SEGMENT S3 (POTENTIAL).
 FT TRANSMEM 495 515 SEGMENT S4 (POTENTIAL).
 FT TRANSMEM 516 532 SEGMENT S5 (POTENTIAL).
 FT DOMAIN 533 549 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 550 570 SEGMENT S6 (POTENTIAL).
 FT TRANSMEM 571 615 SEGMENT S7 (POTENTIAL).
 FT TRANSMEM 616 662 SEGMENT S8 (POTENTIAL).
 FT DOMAIN 663 1195 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 41 70 PAS.
 FT DOMAIN 92 144 PAC.
 FT NP_BIND 745 862 CNMP.
 FT CARBOHYD 600 600 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SO SEQUENCE 1195 AA; 134900 MW; BA24C54B86C5B7 CRC64;

Query Match 54.5%; Score 3312.5; DB 1; Length 1195;
 Best Local Similarity 57.5%; Pred. No. 4.2e-164;
 Matches 724; Conservative 114; Mismatches 251; Indels 171; Gaps 31;

QY 1 MPVRGHAAPONTFLDTIRKEGOSRKTITANAVENCAVTCNDGCELCGYSRAPYM 60
 1 MPVRGHAAPONTFLDTIRKEGOSRKTITANAVENCAVTCNDGCELCGYSRAPYM 60
 DB 1 MPVRGHAAPONTFLDTIRKEGOSRKTITANAVENCAVTCNDGCELCGYSRAPYM 60
 QY 61 QRPCTCDLHGPRTORRAAQIAQALLGAERKVEIAFYRKDSCFCFLVAVVVKNEGD 120
 61 QRPCTCDLHGPRTORRAAQIAQALLGAERKVEIAFYRKDSCFCFLVAVVVKNEGD 120
 DB 61 QRPCTCDLHGPRTORRAAQIAQALLGAERKVEIAFYRKDSCFCFLVAVVVKNEGD 120
 QY 121 AVTFILNFEYVMEKDMGSPAHDTHRGPTSMALPGRATFRKLPLALLARRESY 180
 121 AVTFILNFEYVMEKDMGSPAHDTHRGPTSMALPGRATFRKLPLALLARRESY 180
 DB 121 AVTFILNFEYVMEKDMGSPAHDTHRGPTSMALPGRATFRKLPLALLARRESY 180
 QY 121 VAMFIIINFEVETDEDNAASP-----ERVNPILPVKSVNRKLFGRKFGVLVYRKOSL 175
 121 VAMFIIINFEVETDEDNAASP-----ERVNPILPVKSVNRKLFGRKFGVLVYRKOSL 175
 DB 121 VAMFIIINFEVETDEDNAASP-----ERVNPILPVKSVNRKLFGRKFGVLVYRKOSL 175
 QY 181 RSGAGAGAGAVGVVDVLPAPSSSESLADEVTAMDNVAGHPAEEERATVPG-- 238
 181 RSGAGAGAGAVGVVDVLPAPSSSESLADEVTAMDNVAGHPAEEERATVPG-- 238
 DB 176 PQED-----PDVVYID-----SKHSDSVAMKHKFKSTKSCSPSEADDTKALIQPSQC 225
 176 PQED-----PDVVYID-----SKHSDSVAMKHKFKSTKSCSPSEADDTKALIQPSQC 225
 QY 239 SPPRSAPGQL---PSPRAHSLNDASGSSCSLARTSRSCASVYRASSADIEAMRAG 294
 239 SPPRSAPGQL---PSPRAHSLNDASGSSCSLARTSRSCASVYRASSADIEAMRAG 294
 DB 226 SPLVINSIGPLDSSPKRQMDRLPYDMLQSSQLTHSRRESLCSIRRASSVHDIEGF--N 283
 226 SPLVINSIGPLDSSPKRQMDRLPYDMLQSSQLTHSRRESLCSIRRASSVHDIEGF--N 283
 QY 295 VLPP---PPRIAS-----TGAMHPLRSGILNSTSDSLVRYRISKIPQITLNFVLK 344
 295 VLPP---PPRIAS-----TGAMHPLRSGILNSTSDSLVRYRISKIPQITLNFVLK 344
 DB 284 VHPNIRFDRHASDNGRNVKGPNNHKKSSLLGSSDSNLMKXSTINKITPOLTINFSPDK 343
 284 VHPNIRFDRHASDNGRNVKGPNNHKKSSLLGSSDSNLMKXSTINKITPOLTINFSPDK 343
 QY 345 GDPPLAS-PTSDRETIAPKIKERTHNVTETQVLSLCAADVLPKTIQAPRIHMTIAY 403
 345 GDPPLAS-PTSDRETIAPKIKERTHNVTETQVLSLCAADVLPKTIQAPRIHMTIAY 403
 DB 344 TEKNTSPSSSDKTIIAKVKERTHNVTETQVLSLCAADVLPKTIQAPRIHMTIAY 403
 344 TEKNTSPSSSDKTIIAKVKERTHNVTETQVLSLCAADVLPKTIQAPRIHMTIAY 403
 QY 404 SPKAVMDWLLLVITYAVETPYSAFLKETEGBPPATECGYACOPLAVALVDLIDMF 463
 404 SPKAVMDWLLLVITYAVETPYSAFLKETEGBPPATECGYACOPLAVALVDLIDMF 463
 DB 404 SPKAVMDWLLLVITYAVETPYSAFLKETEGBPPATECGYACOPLAVALVDLIDMF 462
 404 SPKAVMDWLLLVITYAVETPYSAFLKETEGBPPATECGYACOPLAVALVDLIDMF 462
 QY 464 IVDLLINFRITYVANNEVSHPGRIAVHYFKGMFLDMVAALPDDLIIFSGSSEE--L 520
 464 IVDLLINFRITYVANNEVSHPGRIAVHYFKGMFLDMVAALPDDLIIFSGSSEE--L 520
 DB 463 IVDLLINFRITYVANNEVSHPGRIAVHYFKGMFLDMVAALPDDLIIFSGSSEE--L 522
 463 IVDLLINFRITYVANNEVSHPGRIAVHYFKGMFLDMVAALPDDLIIFSGSSEE--L 522
 QY 521 IGLTKTARLRLVYARKLDRYSEYGAVLFLMCTFALLIHMALCIYUAGNMOPMD 580
 521 IGLTKTARLRLVYARKLDRYSEYGAVLFLMCTFALLIHMALCIYUAGNMOPMD 580
 DB 523 IGLTKTARLRLVYARKLDRYSEYGAVLFLMCTFALLIHMALCIYUAGNERPLVT 582
 523 IGLTKTARLRLVYARKLDRYSEYGAVLFLMCTFALLIHMALCIYUAGNERPLVT 582
 QY 581 SRIGLHNLGDOIGKRPYN-SSGLGSPSFKDKYVTALEYTFESSLSVSGVNSPNTNSKI 639
 581 SRIGLHNLGDOIGKRPYN-SSGLGSPSFKDKYVTALEYTFESSLSVSGVNSPNTNSKI 639
 DB 583 DKIGMDSLGQIKRNVDSOSSGSPSFKDKYVTALEYTFESSLSVSGVNSPNTNSKI 642
 583 DKIGMDSLGQIKRNVDSOSSGSPSFKDKYVTALEYTFESSLSVSGVNSPNTNSKI 642
 QY 640 FSICVMLIGSLMYSIRGNVSAIIQRLYSGTARHTOMLRYREFIRHQIDNPLRQLEE 699
 640 FSICVMLIGSLMYSIRGNVSAIIQRLYSGTARHTOMLRYREFIRHQIDNPLRQLEE 699
 DB 643 FSICVMLIGSLMYSIRGNVSAIIQRLYSGTARHTOMLRYREFIRHQIDNPLRQLEE 702
 643 FSICVMLIGSLMYSIRGNVSAIIQRLYSGTARHTOMLRYREFIRHQIDNPLRQLEE 702
 QY 700 YFOHAMSTNGIDMNAVLYKGFPECLQADICHLNRSLLQHKPRGAKGGLRALAMFK 759
 700 YFOHAMSTNGIDMNAVLYKGFPECLQADICHLNRSLLQHKPRGAKGGLRALAMFK 759
 DB 703 YFOHAMSTNGIDMNAVLYKGFPECLQADICHLNRSLLQHKPRGAKGGLRALAMFK 762
 703 YFOHAMSTNGIDMNAVLYKGFPECLQADICHLNRSLLQHKPRGAKGGLRALAMFK 762

QY 760 TTHAPPGDTLVHAGDILLTALYFISRGSTIELRGVVAALLKNDIFGEPLNLVAPRKSN 819
 760 TTHAPPGDTLVHAGDILLTALYFISRGSTIELRGVVAALLKNDIFGEPLNLVAPRKSN 819
 DB 763 TTHAPPGDTLVHAGDILLTALYFISRGSTIELRGVVAALLKNDIFGEPLNLVAPRKSN 822
 763 TTHAPPGDTLVHAGDILLTALYFISRGSTIELRGVVAALLKNDIFGEPLNLVAPRKSN 822
 QY 820 GDVALYLYCDLHKHRRDLEVLDMYPEFSHFSSLEITFNLDNTNIPSPSTELG 879
 820 GDVALYLYCDLHKHRRDLEVLDMYPEFSHFSSLEITFNLDNTNIPSPSTELG 879
 DB 823 ADVALYLYCDLHKHRRDLEVLDMYPEFSHFSLFNLDELNLHRES--AKSOSINDEG 880
 823 ADVALYLYCDLHKHRRDLEVLDMYPEFSHFSLFNLDELNLHRES--AKSOSINDEG 880
 QY 880 GFSORRKRKLSFRRTTOKDPOGEVSALGGRAGAPSS-----GRPGGWGSPSSGP 935
 880 GFSORRKRKLSFRRTTOKDPOGEVSALGGRAGAPSS-----GRPGGWGSPSSGP 935
 DB 881 DTCLRRRLRSFESGDKDFKENSAN-----DADDSTTIRYOSKKNHFEKRS 933
 881 DTCLRRRLRSFESGDKDFKENSAN-----DADDSTTIRYOSKKNHFEKRS 933
 QY 936 SSPSSSEDEGP-----GRSSPLRLVPPSSPPGE---PPG---EPLMEDCKSSDT 983
 936 SSPSSSEDEGP-----GRSSPLRLVPPSSPPGE---PPG---EPLMEDCKSSDT 983
 DB 934 SSFSSIDDEQKPLFLGTVDSTPMVYKAS--RHGGEAPPSGRLHMDKRSCKDITDT 991
 934 SSFSSIDDEQKPLFLGTVDSTPMVYKAS--RHGGEAPPSGRLHMDKRSCKDITDT 991
 QY 984 -----CNPLSGAFSGVSNFSGWGRGROYELRCRAPPTSLNIPSS 1029
 984 -----CNPLSGAFSGVSNFSGWGRGROYELRCRAPPTSLNIPSS 1029
 DB 992 HSWERENARAQPEBCSP-----SGIQR--AAWGISE-----TESDLYT 1027
 992 HSWERENARAQPEBCSP-----SGIQR--AAWGISE-----TESDLYT 1027
 QY 1030 PGRPRGPDVESRLDALORLRLFTRLSADATVLIOLLOMTLVPPAYSAVT----- 1082
 1030 PGRPRGPDVESRLDALORLRLFTRLSADATVLIOLLOMTLVPPAYSAVT----- 1082
 DB 1028 -----GEVEGRDLLEQLRLKLSQMTTDIQAIIQLQKQTYVPAYSMVYAGAEYQR 1081
 1028 -----GEVEGRDLLEQLRLKLSQMTTDIQAIIQLQKQTYVPAYSMVYAGAEYQR 1081
 QY 1083 -----TPGP-----GPTSTP-----LLPVSPLPLTLDLSIQ 1110
 1083 -----TPGP-----GPTSTP-----LLPVSPLPLTLDLSIQ 1110
 DB 1082 PILRLFTLRSHRASIKTDRSFSQCEFLDLKSKLSKESLSGKRLTASSEDNITS 1141
 1082 PILRLFTLRSHRASIKTDRSFSQCEFLDLKSKLSKESLSGKRLTASSEDNITS 1141
 QY 1111 V--SQFACCELPPGA-----PELPQEGPTRLSLGOLGALTSQPLHGHGSPDG 1158
 1111 V--SQFACCELPPGA-----PELPQEGPTRLSLGOLGALTSQPLHGHGSPDG 1158
 DB 1142 LKQDSDASSSELDPORKSYAPIRHPSLPDSS-----LSTVGILG-----LHRHVSOPG 1191
 1142 LKQDSDASSSELDPORKSYAPIRHPSLPDSS-----LSTVGILG-----LHRHVSOPG 1191

RESULT 7
 KCH7_MOUSE
 ID KCH7_MOUSE STRAND: PRT: 1195 AA.
 AC 09ER47;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Potassium voltage-gated channel subfamily H member 7 (Ether-a-go-go related gene potassium channel 3) (Ether-a-go-go related protein 3).
 DE (Eag-related protein 3).
 GN KCNH7 OR ERG3.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RA Arcangeli A.;
 RT "Erg genes expression during development of mouse embryos."
 RL Submitted (Oct-2000) to the EMBL/Genbank/DBJ databases.
 RC -i- FUNCTION: Pore-forming (alpha) subunit of voltage-gated potassium channel (by similarity). Channel properties may be modulated by cAMP and subunit assembly.
 CC -i- SUBUNIT: The potassium channel is probably composed of a homo- or heterotetrameric complex of pore-forming alpha subunits that can associate with modulating beta subunits. Heteromultimer with KCNH2/ERG1 and KCNH6/ERG2 (by similarity).
 CC -i- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -i- DOMAIN: The segment S4 is probably the voltage-sensor and is characterized by a series of positively charged amino acids at every third position.
 CC -i- SIMILARITY: Belongs to the potassium channel family. H (Eag) subfamily.
 CC -i- SIMILARITY: Contains 1 cyclic nucleotide-binding domain.
 CC -i- SIMILARITY: Contains 1 PAS (PER-ARNT-SIM) dimerization domain.
 CC -i- SIMILARITY: Contains 1 PAS-associated C-terminal (PAC) domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL: AJ291608; CAC14797.1; -
DR InterPro: IPR000595; cNMP_binding.
DR InterPro: IPR003967; Erg_channel.
DR InterPro: IPR005821; Ion_trans.
DR InterPro: IPR001622; K_channel_pore.
DR InterPro: IPR005820; M_channel_nlg.
DR InterPro: IPR001610; PAC.
DR InterPro: IPR000700; PAS-assoc.C.
DR InterPro: IPR000014; PAS_domain.
DR Pfam: PF000027; cNMP_binding.1.
DR Pfam: PF00520; Ion_trans.1.
DR Pfam: PF00785; PAC.1.
DR SMART: SM00100; cNMP.1.
DR PROSITE: PS50042; cNMP_BINDING_3.1.
DR PROSITE: PS50113; PAC; FALSE_NEG.
DR PROSITE: PS50112; PAS; FALSE_NEG.
KW Transport; Ion transport; Ionic channel; Voltage-gated channel;
KW Potassium channel; Potassium; Potassium transport; Transmembrane;
KW Glycoprotein; Multigene family.
FT DOMAIN 1 412 CYTOPLASMIC (POTENTIAL).
FT TRANSSEM 413 433 SEGMENT S1 (POTENTIAL).
FT TRANSSEM 450 470 SEGMENT S2 (POTENTIAL).
FT DOMAIN 471 494 CYTOPLASMIC (POTENTIAL).
FT TRANSSEM 495 515 SEGMENT S3 (POTENTIAL).
FT TRANSSEM 522 542 SEGMENT S4 (POTENTIAL).
FT DOMAIN 543 549 CYTOPLASMIC (POTENTIAL).
FT TRANSSEM 550 570 SEGMENT S5 (POTENTIAL).
FT DOMAIN 615 632 SEGMENT H5 (PORE-FORMING) (POTENTIAL).
FT TRANSSEM 642 662 SEGMENT S6 (POTENTIAL).
FT DOMAIN 663 1195 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 41 70 PAS.
FT NP_BIND 92 144 cNMP.
FT CARBOHYD 745 862 N-LINKED (GLCNAC...) (POTENTIAL).
SQ SEQUENCE 1195 AA; 135026 MW; 8DCB9BA0580FEC2 CRC64;

Query Match 54.0%; Score 3283.5; DB 1; Length 1195;

Best Local Similarity 56.4%; Pred. No. 1.3e-162; Mismatches 253; Indels 181; Gaps 30;

Matches 714; Conservative 117; Mismatches 253; Indels 181; Gaps 30;

QY 1 MPVARGHVAPONTFLDITIRKFEQSGSRKFTIANARVENCAYICNDGFCICGYSRAEVM 60
Db 1 MPVARGHVAPONTFLDITIRKFEQSGSRKFTIANARVENCAYICNDGFCICGYSRAEVM 60

QY 61 QRPCTCFELHGRFTRORAAQOIAALLGAERKVEIAFYRKDSCFCICLVDPVYKVEDG 120
Db 61 QRPCTCFELHGRFTRORAAQOIAALLGAERKVEIAFYRKDSCFCICLVDPVYKVEDG 120

QY 121 AVINFTILFEVNEKDWGSPADTNHGRPTSLVAGRAKTFELKIPALLATLARSSV 180
Db 121 AVINFTILFEVNEKDWGSPADTNHGRPTSLVAGRAKTFELKIPALLATLARSSV 180

QY 121 VAMMFTILFEVYDEENAAFP-----ERVNPILPVKTVNKRKLGFKFGRLVLYTRKOSL 175
Db 121 VAMMFTILFEVYDEENAAFP-----ERVNPILPVKTVNKRKLGFKFGRLVLYTRKOSL 175

QY 181 RSGGAGAGAPGAVVVDVLTLPAPASSESLALDEVTAMDNVAGLGAERKRALVGP-- 238
Db 181 RSGGAGAGAPGAVVVDVLTLPAPASSESLALDEVTAMDNVAGLGAERKRALVGP-- 238

QY 176 PQED-----PDVVVID-----SKHSDSVAMKHFKSPTEKSCSPSADDTKALIQSSQC 225
Db 176 PQED-----PDVVVID-----SKHSDSVAMKHFKSPTEKSCSPSADDTKALIQSSQC 225

QY 239 SPERSARFGL-----PSRAISLNPDASGSSCSIAKTRSRSCASVPRASSADITEAMRAG 294
Db 239 SPERSARFGL-----PSRAISLNPDASGSSCSIAKTRSRSCASVPRASSADITEAMRAG 294

QY 226 SPLVNISGPLDHSPPKRWMDRLYDMLQSSQLTHSRKSRSLCSIRASSVHDIIEGF--S 283
Db 226 SPLVNISGPLDHSPPKRWMDRLYDMLQSSQLTHSRKSRSLCSIRASSVHDIIEGF--S 283

QY 295 VLPP-----PPRHAS-----TGAMHPLRSGLSLNTSDPVRRTTISKIQTILNFVDLK 344
Db 295 VLPP-----PPRHAS-----TGAMHPLRSGLSLNTSDPVRRTTISKIQTILNFVDLK 344

QY 284 VHKKNIFRDRHASEDNRRANKGPPNHRKSSLSLGSSTSNLNKSTIKIKIQLTILNPSDK 343
Db 284 VHKKNIFRDRHASEDNRRANKGPPNHRKSSLSLGSSTSNLNKSTIKIKIQLTILNPSDK 343

QY 345 GDFPLAS-PTSDBEIIAPKIKERTHNVTEKVTVLSGADVDPYKLAQAPRIHRWTILHY 403
Db 345 GDFPLAS-PTSDBEIIAPKIKERTHNVTEKVTVLSGADVDPYKLAQAPRIHRWTILHY 403

Db 344 TEKNTSPSPSDKTIIPAKVERTHNTEKVTVLSGADVDPYKLAQAPRIHRWTILHY 403

QY 404 SPFAVMDWILLIVITATFTPSAFLKEPDEGPRATCEGTAQCPPLAVVDIYIMF 463

Db 404 SPFAVMDWILLIVITATFTPSAFLKEPDEGPRATCEGTAQCPPLAVVDIYIMF 463

QY 464 IVDILINFRRTTYVANEVSSHPRIAVHFYFKGFELDMVAALPFDDLIFGSGSEE--L 520

Db 463 IVDILINFRRTTYVANEVSSHPRIAVHFYFKGFELDMVAALPFDDLIFGSGSEE--L 520

QY 521 IGLKTARLLRLVAVARKLDRESEYGAVALFELMCTPALAHMLACIWAIGNNEOPDM 580

Db 523 IGLKTARLLRLVAVARKLDRESEYGAVALFELMCTPALAHMLACIWAIGNNEOPDM 580

QY 581 SRIGMLNLDQIGKFPIN-SSGLGSPSIRKQYATALFTSSLSVGFVNSPPTNSEKI 639

Db 583 DKIGMLDSTQIGKRYNDSSSGPIKQYVALYFTSSLSVGFVNSPPTNSEKI 642

QY 640 FSICVMLIGSLMVASIFGNVSAIQRLYSGTARYHTQMLVREFIRFHQINPLRORLEE 699

Db 643 FSICVMLIGSLMVASIFGNVSAIQRLYSGTARYHTQMLVREFIRFHQINPLRORLEE 702

QY 700 YFOHAWSTYNGIDMNAVALKGFPEECLOADICILNRSILQHKPREGATKGLRALAMKFK 759

Db 703 YFOHAWSTYNGIDMNAVALKGFPEECLOADICILNRSILQHKPREGATKGLRALAMKFK 762

QY 760 TTAPRPEDTLVHAGDILITALYFSRGSIEILKGDVVAALIGKNDISEPILNLYARPKSN 819

Db 763 TTAPRPEDTLVHAGDILITALYFSRGSIEILKGDVVAALIGKNDISEPILNLYARPKSN 822

QY 820 GDVVALYCDLKHHRDLEVDMPEDSHFESSLEITFNLDTMIGSPGSTLEG 879

Db 823 ADVRALYCDLKHHRDLEVDMPEDSHFESSLEITFNLDTMIGSPGSTLEG 880

QY 880 GFGRKRRKLSFRRTDKTEQDEVSALGPRAGAPSS-----RGRPGRPWGESPPSGP 935

Db 881 GFGRKRRKLSFRRTDKTEQDEVSALGPRAGAPSS-----RGRPGRPWGESPPSGP 933

QY 936 SSPSSDEDEGRKSSSLRL-VPPSSPR-----PQEPGGERPLMDECKSS 981

Db 934 SSPSSDEDEGRKSSSLRL-VPPSSPR-----PQEPGGERPLMDECKSS 989

QY 982 DT-----CNPLSGAFSGVSNIFSEWGDGRGOYOLPRCPAPTSILNIPL 1027

Db 990 DT-----CNPLSGAFSGVSNIFSEWGDGRGOYOLPRCPAPTSILNIPL 1027

QY 1028 \$SPGRPRGDVESRLDALQRLNLETRLSADMAVTVLQLRQMTLVPPAVSAVTPPG 1087

Db 1028 \$SPGRPRGDVESRLDALQRLNLETRLSADMAVTVLQLRQMTLVPPAVSAVTPPG 1076

QY 1088 PTSTSPLL-----PVPLP-----TLTL 1105

Db 1077 PTSTSPLL-----PVPLP-----TLTL 1105

QY 1106 DSLQOV--SOFMACEELPPGA-----PELPQGPTRRLSLPQLGALTSQPLHRH 1153

Db 1137 DSLQOV--SOFMACEELPPGA-----PELPQGPTRRLSLPQLGALTSQPLHRH 1186

QY 1154 GSDPG 1158

Db 1187 GSDPG 1191

RESULT 8

KCH7_HUMAN STANDARD; PRT; 1196 AA.

AC ONS40;

DT 28-FEB-2003 (Rel. 41, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Potassium voltage-gated channel subfamily H member 7 (Ether-a-go-go related gene potassium channel 3) (HERG-3) (Ether-a-go-go related

DE protein 3) (Eag related protein 3).
 CN KCMH7 OR ERG3.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 RN NCBI_TaxID=9606;
 [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RA Ganetzky B., Titus S.A.;
 RT "Polynucleotides encoding herg-3 potassium channel."
 RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.
 CC -1 FUNCTION: Pore-forming (alpha) subunit of voltage-gated potassium
 CC channel. Channel properties may be modulated by cAMP and subunit
 CC assembly.
 CC -1 SUBUNIT: The potassium channel is probably composed of a homo- or
 CC heterotetrameric complex of pore-forming alpha subunits that can
 CC associate with modulating beta subunits. Heteromultimer with
 CC KCMH2/ERG1 and KCMH6/ERG2 (By similarity).
 CC -1 SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1 DOMAIN: The segment S4 is probably the voltage-sensor and is
 CC characterized by a series of positively charged amino acids at
 CC every third position.
 CC -1 SIMILARITY: Belongs to the potassium channel family. H (Eag)
 CC subfamily.
 CC -1 SIMILARITY: Contains 1 cyclic nucleotide-binding domain.
 CC -1 SIMILARITY: Contains 1 PAS (PBR-ARNT-SIM) dimerization domain.
 CC -1 SIMILARITY: Contains 1 PAS-associated C-terminal (PAC) domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL: AF032897; AAD01946.1; -
 DR Genbank: HGNC:18863; KCMH7.
 DR InterPro: IPR000595; CNMP_binding.
 DR InterPro: IPR003967; Erg_channel.
 DR InterPro: IPR005821; Ion_trans.
 DR InterPro: IPR001622; K+channel_pore.
 DR InterPro: IPR005820; M+channel_nlg.
 DR InterPro: IPR001610; PAC.
 DR InterPro: IPR000700; PAS-assoc_C.
 DR InterPro: IPR000014; PAS_domain.
 DR Pfam: PF00027; CNMP_binding; 1.
 DR Pfam: PF00520; Ion_trans; 1.
 DR Pfam: PF00785; PAC; 1.
 DR SMART: SM00100; CNMP; 1.
 DR SMART: SM00086; PAC; 1.
 DR SMART: SM00091; PAS; 1.
 DR PROSITE: PS50042; CNMP_BINDING_3; 1.
 DR PROSITE: PS50113; PAC; FALSE_NEG.
 DR PROSITE: PS50112; PAS; FALSE_NEG.
 KW Transport; ion transport; Ionic channel; Voltage-gated channel;
 KW Potassium channel; Potassium; Potassium transport; Transmembrane;
 KW Glycoprotein; Multigene family.
 FT DOMAIN 1 412
 FT TRANSEM 413 433 SEGMENT S1 (POTENTIAL).
 FT TRANSEM 450 470 SEGMENT S2 (POTENTIAL).
 FT DOMAIN 471 494 CYTOPLASMIC (POTENTIAL).
 FT TRANSEM 495 515 SEGMENT S3 (POTENTIAL).
 FT TRANSEM 522 542 SEGMENT S4 (POTENTIAL).
 FT DOMAIN 543 549 CYTOPLASMIC (POTENTIAL).
 FT TRANSEM 550 570 SEGMENT S5 (POTENTIAL).
 FT DOMAIN 615 635 SEGMENT S6 (POTENTIAL).
 FT TRANSEM 642 662 SEGMENT S6 (POTENTIAL).
 FT DOMAIN 663 1196 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 92 144 PAS.
 FT NP_BIND 745 862 CNMP.

FT CAROBYD 600 600 N-LINKED (GLCNAC...) (POTENTIAL).
 SQ SEQUENCE 1196 AA; 135012 MW; 7CE10C31A76D4FE CRC64;
 Query Match 53.5%; Score 3255; DB 1; Length 1196;
 Best Local Similarity 54.8%; Pred. No. 4e-161;
 Matches 700; Conservative 126; Mismatches 246; Indels 206; Gaps 25;
 1 MPVRGHVAPONTFLDTIIRKREGOSRKFTIARAVENCAYLYCNDGFCGLGYRAEVM 60
 1 MPVRGHVAPONTFLDTIIRKREGONKFTIARAVQCALIYICNDGFCGLGYRPPVM 60
 61 QRPCTCDLHGRPRDORRAAQAQALLGAERKEVIATYRKDSCFCLVDVVPYKKNDE 120
 61 QKPCDCDLHGRPRKHDAQIAQALLGSEERKEVITYYHNKGSFTFNTNTHIIPVKNDE 120
 121 AVIMEFLFEVYMEKDWGSPAHDTNHRGPTSMALPGRATFRLKLPALLALARESSV 180
 121 VAMFPIFEVYTDENAKATP-----EVNPIPLKTYNKKFPGFKFGLVLYRKOSL 175
 181 RSGGAGAGAAVAVVDLTPAAPSSSLALDEVYAMDNHVAIGLPAEERRALVGP-- 238
 176 POED-----PDVVYID----SKHSDSVAMKHKSPFKSCSPSEADPTKALIQPKC 225
 239 SPPRSAPQOL-----PSPRAHSLNDPASGSSCLARTRESCASYRRASSADIEAMAG 294
 226 SPLVNIQPLDHSSEPRKQMDRLPYDMLQSSQLSHRSRESLCSIRRASSVHIDEG--G 283
 295 VLPP-----PPIIAS-----TGAMHPLRSGLNTSDSDLVRYRTISKIPOTLNFVYLK 344
 284 VHPKNIFPRDIASENGNVAAGPFRNHKISSLLGSTDENLKYSTINKIPOLITLNFSEVK 343
 345 GDPFLAS-PTSDREIIPAKIKERTHNTEKYTVYLSGADVLPEYKIQAPRIHMTILHY 403
 344 TEKKNSPPSSDSTIIPAKVDRHNTEKYTVYLSGADVLPEYKIQAPRIHMTILHY 403
 404 SPPKAVMDMLLLIYIYAVFTPYSAALFKETEGPATCGYACQPLAVVDLVIDMF 463
 404 SPPKAVMDMLLLIYIYAVFTPYSAALFKETEGPATCGYACQPLAVVDLVIDMF 462
 464 IVDILINERTYVYANEEVSHPRIAVHYKGFLLIDMAVAFDILLFSGSSEE---L 520
 463 IIDILINERTYVYANEEVSDPAKIAHYKGFLLIDMAVAFDILLFSGSDETTTL 522
 521 IGLKTAFLRLVYARKIDRYSEYGAAYFLMCTALIAHMLACIYATGNEDQPHMD 580
 523 IGLKTAFLRLVYARKIDRYSEYGAAYFLMCTALIAHMLACIYATGNEDQPHMD 582
 581 SRIGMHLNLGQIGKRPYK-SSGLGSPSTKQKYYVALYFTFSSLSYVGNGVSPNTNSKI 639
 583 DKIGMLDLSLGGQIGKRYNDSGSSGSPSTKDYVALYFTFSSLSYVGNGVSPNTNSKI 642
 640 FSIQVMLGSLMAYASIFGNYSATIIQRLYSGTARYHTQMLRVREFIRFQIIPNRLQLEE 699
 643 FSIQVMLGSLMAYASIFGNYSATIIQRLYSGTARYHTQMLRVREFIRFQIIPNRLQLEE 702
 700 YFOHAWSYTYNGIDMNAVLYKGFPECLQADICHLNLSLLOHCKPREGATKGLALAMKFK 759
 703 YFOHAWSYTYNGIDMNAVLYKGFPECLQADICHLNLSLLOHCKPREGATKGLALAMKFK 762
 760 TTHAPRGDTLVHADDLTALYFIRGSEIILRGVYVALILGKNIDFEPPLNLYARPKSN 819
 763 TTHALQGTTLVHCDVLTALYFIRGSEIILRGVYVALILGKNIDFEPPLNLYARPKSN 822
 820 GDVALYLYCDLHKIHRDILLEVMYEPFSSLEITFNLLDNTN---IPSPSTE 876
 823 ADVALYLYCDLHKIHRDILLEVMYEPFSSLEITFNLLDNTN---IPSPSTE 882
 877 LEGFSRQRRKRLSFRRTDKDTPQEPVSALGPRAGAGPSSRGPRGPGWGESPPSSGPS 936
 883 SEGQNCRLRRRKLSFESGEKE-----EGPGRSSPLRLVFPSSPPGPGPGLMEDCEKSSDT 983
 937 SPESSED-----EGPGRSSPLRLVFPSSPPGPGPGLMEDCEKSSDT 983


```

FT TRANSMEM 341 361 SEGMENT S3 (POTENTIAL).
FT TRANSMEM 371 391 SEGMENT S4 (POTENTIAL).
FT DOMAIN 392 398 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 399 419 SEGMENT S5 (POTENTIAL).
FT DOMAIN 464 484 SEGMENT H5 (PORE-FORMING) (POTENTIAL).
FT TRANSMEM 491 511 SEGMENT S6 (POTENTIAL).
FT DOMAIN 512 519 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 41 70 PAC.
FT DOMAIN 92 144 PAC.
FT NP_BIND 594 711 CNMP.
FT VASPLIC 419 472 MYAIGVERPYLEHKGICGLDLSGVOLGKRYNGSDPASPSV
FT ODKYVYALXPTFS -> C (in isoform 2).
FT VASPLIC 745 780 /FTid=VSP_000977.
FT VASPLIC 501 502 Missing (in isoform 2).
FT VASPLIC 503 502 /FTid=VSP_000978.
FT VASPLIC 503 994 /FTid=VSP_000979.
FT VASPLIC 503 994 Missing (in isoform 3).
FT CONFLICT 925 925 T -> M (in REF. 2; BAC03764).
FT CONFLICT 963 963 F -> L (in REF. 2; BAC03764).
SQ SEQUENCE 994 AA; 109924 MW; BE9ECB349A798576 CRC64;

Query Match 50.1%; Score 3043.5; DB: 1; Length 994;
Best Local Similarity 54.2%; Pred. No. 2.8e-150;
Matches 670; Conservative 85; Mismatches 153; Indels 329; Gaps 32;

1 MPVRGHAAPONTFTDITIRKFEQSRKFIITANAVENCAYVNDGFCGLGYSRALVM 60
1 MPVRGHAAPONTYDITIRKFEQSRKFIITANOMEMCAIITYCDGFCGLGYSRALVM 60
61 QRPCTCDFLHGRPRQRRRAAQIAQALLGAERKVEIAFYRKDGSCFLVDVVPVKNEDG 120
61 QRPCTCDFLHGRPRQRRRAAQIAQALLGAERKVEIAFYRKDGSCFLVDVVPVKNEDG 120
61 QRPCTCDFLHGRPRQRRRAAQIAQALLGAERKVEIAFYRKDGSCFLVDVVPVKNEDG 120
61 QRPCTCDFLHGRPRQRRRAAQIAQALLGAERKVEIAFYRKDGSCFLVDVVPVKNEDG 120
121 AVIMFLMEFYVMKEDMGSPAHDTHNHRGPTSMAPRAKTRFLKLPALLA-LTARESS 179
121 AVIMFLMEFYVMKEDMGSPAHDTHNHRGPTSMAPRAKTRFLKLPALLA-LTARESS 179
121 AVIMFLMEFYVMKEDMGSPAHDTHNHRGPTSMAPRAKTRFLKLPALLA-LTARESS 179
121 AVIMFLMEFYVMKEDMGSPAHDTHNHRGPTSMAPRAKTRFLKLPALLA-LTARESS 179
180 VR-----SGAGAGAGARVAVVDLTPAAPSSESLALDEVYAMDNHVGAGPAEERRA 233
180 VR-----SGAGAGAGARVAVVDLTPAAPSSESLALDEVYAMDNHVGAGPAEERRA 233
146 QRLSQSFLGSESGHGRG----- 164
234 LVPGSPRSPAPGOLPSPRAHSLNDASGSSCSLTARTSRSCASVRRASSADIEAMRA 293
234 LVPGSPRSPAPGOLPSPRAHSLNDASGSSCSLTARTSRSCASVRRASSADIEAMRA 293
165 --GPG-----PG----- 169
294 GVLPRRRHASTGAMHPLRSGLNSTSDSLVRYRTKIQITLNFYDLKGDPLASPT 353
294 GVLPRRRHASTGAMHPLRSGLNSTSDSLVRYRTKIQITLNFYDLKGDPLASPT 353
170 -----TGR-----GKYPTISQIDPQPTLNFVEFNLEKRRSSST 201
354 SDRELIAP-KIKERTHNTKVTQVLSGADVLPREYKQADAPRIHMTILHYSPEKAVDM 412
354 SDRELIAP-KIKERTHNTKVTQVLSGADVLPREYKQADAPRIHMTILHYSPEKAVDM 412
202 TEIETIAHAKYVERQNTKVTQVLSGADVLPREYKQADAPRIHMTILHYSPEKAVDM 261
202 TEIETIAHAKYVERQNTKVTQVLSGADVLPREYKQADAPRIHMTILHYSPEKAVDM 261
413 LILLVITYAVFTPSAFLKETEPRATECGAACOPPLAVVDLIVIMFVLDILINR 472
413 LILLVITYAVFTPSAFLKETEPRATECGAACOPPLAVVDLIVIMFVLDILINR 472
413 LILLVITYAVFTPSAFLKETEPRATECGAACOPPLAVVDLIVIMFVLDILINR 472
413 LILLVITYAVFTPSAFLKETEPRATECGAACOPPLAVVDLIVIMFVLDILINR 472
262 LILLVITYAVFTPSAFLKETEPRATECGAACOPPLAVVDLIVIMFVLDILINR 320
262 LILLVITYAVFTPSAFLKETEPRATECGAACOPPLAVVDLIVIMFVLDILINR 320
473 TTYVANEEVYSHPRIVHYFKGFLIDMVAALPFDLLIRTSODETTTLGLKTLARL 529
473 TTYVANEEVYSHPRIVHYFKGFLIDMVAALPFDLLIRTSODETTTLGLKTLARL 529
473 TTYVANEEVYSHPRIVHYFKGFLIDMVAALPFDLLIRTSODETTTLGLKTLARL 529
473 TTYVANEEVYSHPRIVHYFKGFLIDMVAALPFDLLIRTSODETTTLGLKTLARL 529
321 TTYVANEEVYSHPRIVHYFKGFLIDMVAALPFDLLIRTSODETTTLGLKTLARL 380
321 TTYVANEEVYSHPRIVHYFKGFLIDMVAALPFDLLIRTSODETTTLGLKTLARL 380
530 LRLVAVARKLDYRSEYGAVALFLMCTPALLIAHMTIACIWMYAGNEREYLEKIKWMDSL 589
530 LRLVAVARKLDYRSEYGAVALFLMCTPALLIAHMTIACIWMYAGNEREYLEKIKWMDSL 589
381 LRLVAVARKLDYRSEYGAVALFLMCTPALLIAHMTIACIWMYAGNEREYLEKIKWMDSL 440
381 LRLVAVARKLDYRSEYGAVALFLMCTPALLIAHMTIACIWMYAGNEREYLEKIKWMDSL 440
590 GDQJGKPYNSG-LGSPSICKKYVYALYFTSSLTSGFGVNSPNTSEKIFPSICVMLIG 648
590 GDQJGKPYNSG-LGSPSICKKYVYALYFTSSLTSGFGVNSPNTSEKIFPSICVMLIG 648
441 GVOJGKRYNSGDPASGSPVODKYVYALYFTSSLTSGFGVNSPNTSEKIFPSICVMLIG 500
441 GVOJGKRYNSGDPASGSPVODKYVYALYFTSSLTSGFGVNSPNTSEKIFPSICVMLIG 500
649 SLMASTIGNSATIORLYSGTARYHOMLVREIRRHOTPNPLROLEEFQHANST 708
649 SLMASTIGNSATIORLYSGTARYHOMLVREIRRHOTPNPLROLEEFQHANST 708

```

```

Db 501 SLMASTIGNSATIORLYSGTARYHOMLVREIRRHOTPNPLROLEEFQHANST 560
Qy 709 NGIDMNAVLKPEPECLQADICLHNRSLLOCKPREGATKGCRLALANKFTTAPPGT 768
Db 561 NGIDMNAVLKPEPECLQADICLHNRSLLOCKPREGATKGCRLALANKFTTAPPGT 620
Qy 769 LVHAGDILYALYFTSRGSEIELRGDYYVAIIIGKNDIFEEPLNTLARPCKSGDVALTYC 828
Db 621 LVHAGDILYALYFTSRGSEIELRGDYYVAIIIGKNDIFEEPLNTLARPCKSGDVALTYC 680
Qy 829 DLHKIRHRLDILEVDMYEPFSDHFWSSLEITFNLDT-----NMIPGSGSTELEGFSR 883
Db 681 DLHKIRHRLDILEVDMYEPFSDHFWSSLEITFNLDT-----NMIPGSGSTELEGFSR 738
Qy 884 QRRKRLSFRRTDKDTEOPGSEVSAAGPAGAGPSSRPGPGPGEPSGSPS--SPSS 941
Db 739 -----LSDNGSGSPHE--LGP-----QFPSKGYSLIGPSQ 767
Qy 942 EDEGPRSSPLRLVYFPSSPPRPGEPGGEPLMEDCKESDTCNPLSGAFSGVSNIFSFW 1001
Db 768 NSMGAG-----PCAPGHDAAPL-----SISDASGLW 795
Qy 1002 GDSRGROYQELPCPCAPPTPILLNPLSSPGRPRGDE-----SRDALOROLNRET 1054
Db 796 PE-----LLQEMP-----PRHSP-OSPOEDPCWPLKLSRLQJLQAOQNRLES 838
Qy 1055 RLSDMATVQLDLQROM-----TLVPAYSAVTPGP-----GPTS 1090
Db 839 RVSSDLSLTLQDLQKPMQGHASYTLAPASNDLALVPIA-SETTSPGRPLPGFLPPAQ 897
Qy 1091 T-----SPLPVPSPPLTLTDSLSQVSQFMACELTLPG-----APLPP 1128
Db 898 TPYSYGDLDCCSPKHNRSPPRMP--HLAVATKTLAPSS--OEOPGLMPLASPLHP 951
Qy 1129 OE-----GPTRRRLSPQLGALTQ--PLHRHSGDPG 1158
Db 952 LEVQGLIGPCFS-SLPEHLSVPRQDLDFQRHSGDPG 987

RESULT 10
KCH6_RAT
ID KCH6_RAT STANDARD: PRT: 950 AA.
AC OS4853.
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DE Potassium voltage-gated channel subfamily H member 6 (Ether-a-go-go related gene potassium channel 2) (Ether-a-go-go related protein 2).
DE (Eag related protein 2).
GN KCH6 OR ERG2.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Cervical ganglion;
RX MEDLINE=98054206; PubMed=939098;
RA Shi W., Wyomere R.S., Wang H.-S., Pan Z., Cohen I.S., McKinnon D.,
RA Dixon J.E.;
RT "Identification of two nervous system-specific members of the erg
RT potassium channel gene family.";
RL J. Neurosci. 17:9423-9432(1997).
RN [2]
RP TISSUE SPECIFICITY.
RX MEDLINE=20183472; PubMed=10718922;
RA Wulfsen I., Hauber H.P., Schiemann D., Bauer C.K., Schwarz J.R.;
RT "Expression of mRNA for voltage-dependent and inward-rectifying K
RT channels in GH3/B6 cells and rat pituitary.";
RL J. Neuroendocrinol. 12:263-272(2000).
RN [3]
RP INTERACTION WITH KCHN2 AND KCHN7, AND MUTAGENESIS OF GLY-480.
RX MEDLINE=21079731; PubMed=11212207;

```

RA Wimmers S., Wulfsen I., Bauer C.K., Schwarz J.R.;
 RT "Erg1, erg3 and erg3 K channel subunits are able to form
 RT heteromultimers.";
 RL Pfugers Arch. 441:450-455(2001).
 RN [4]
 RP TISSUE SPECIFICITY.
 RX MEDLINE=21319165; PubMed=11425889;
 RA Saganich M.J., Machado E., Rudy B.;
 RT "Differential expression of genes encoding subthreshold-operating
 RT voltage-gated K⁺ channels in brain.";
 RL J. Neurosci. 21:4609-4624(2001).
 CC -1- FUNCTION: Pore-forming (alpha) subunit of voltage-gated potassium
 CC channel. Elicits a slowly activating, rectifying current. Channel
 CC properties may be modulated by cAMP and subunit assembly.
 CC -1- SUBUNIT: The potassium channel is probably composed of a homo- or
 CC heterotrimeric complex of pore-forming alpha subunits that can
 CC associate with modulating beta subunits. Heterotrimer with
 CC KCHN2/ERG1 and KCNH7/ERG3.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- TISSUE SPECIFICITY: Highly expressed in celiac and superior
 CC mesenteric ganglia, but not detected in brain or in heart.
 CC Detected at low levels in retina. Also found in pituitary.
 CC According Ref.4 found in olfactory bulb (granular and mitral cell
 CC layers).
 CC -1- DOMAIN: The segment s4 is probably the voltage-sensor and is
 CC characterized by a series of positively charged amino acids at
 CC every third position.
 CC -1- SIMILARITY: Belongs to the potassium channel family. H (Eag)
 CC subfamily.
 CC -1- SIMILARITY: Contains 1 cyclic nucleotide-binding domain.
 CC -1- SIMILARITY: Contains 1 PAS (PER-ARNT-SIM) dimerization domain.
 CC -1- SIMILARITY: Contains 1 PAS-associated C-terminal (PAC) domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL, AF016192; AB94742.1; -
 DR InterPro: IPR000595; CNMP_binding.
 DR InterPro: IPR003967; Erg_channel.
 DR InterPro: IPR005821; Ion_trans.
 DR InterPro: IPR001622; K_channel_pore.
 DR InterPro: IPR005820; M_channel_intg.
 DR InterPro: IPR001610; PAC.
 DR InterPro: IPR000700; PAS_assoc_C.
 DR Pfam: PF000027; CNMP_binding; 1.
 DR Pfam: PF005520; Ion_trans; 1.
 DR Pfam: PF00785; PAC; 1.
 DR SMART: SM00100; CNMP; 1.
 DR SMART: SM00086; PAC; 1.
 DR PROSITE: PS50042; CNMP_BINDING_3; 1.
 DR PROSITE: PS50113; PAC; FALSE_NEG.
 DR PROSITE: PS50112; PAS; FALSE_NEG.
 KM Transport; Ion transport; Ionic channel; Voltage-gated channel;
 KM Potassium channel; Potassium; Potassium transport; Transmembrane;
 KM Multigene family.
 FT DOMAIN 1 261 CYTOPLASMIC (POTENTIAL).
 FT TRANSSEM 262 282 SEGMENT S1 (POTENTIAL).
 FT TRANSSEM 299 319 SEGMENT S2 (POTENTIAL).
 FT DOMAIN 320 340 CYTOPLASMIC (POTENTIAL).
 FT TRANSSEM 341 361 SEGMENT S3 (POTENTIAL).
 FT TRANSSEM 371 391 SEGMENT S4 (POTENTIAL).
 FT TRANSSEM 392 398 CYTOPLASMIC (POTENTIAL).
 FT TRANSSEM 399 419 SEGMENT S5 (POTENTIAL).
 FT DOMAIN 464 484 SEGMENT H5 (PORE-FORMING) (POTENTIAL).
 FT TRANSSEM 491 511 SEGMENT S6 (POTENTIAL).
 FT DOMAIN 512 950 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 41 70 PAS.

FT DOMAIN 92 144 PAC.
 FT NP_BIND 594 711 CNMP.
 FT MUTAGEN 480 480 G->S: DOMINANT NEGATIVE MUTANT; ABOLISHES
 FT SEQUENCE 950 AA: 105705 MW: 455674.8087760 CRG64;
 SO QUERY MATCH 49.1%; Score 2982; DB 1; Length 950;
 Best Local Similarity 52.9%; Pred. No. 4e-147;
 Matches 647; Conservative 87; Mismatches 146; Indels 344; Gaps 25;
 QY 1 MPVRGVAQNTFLDTITIRKFEQSKRTIANKAVENCAYITNDGFCICGSRAEVM 60
 DB 1 MPVRGVAQNTFLDTITIRKFEQSKRTIANKAVENCAYITNDGFCICGSRAEVM 60
 QY 61 QRPCTGFLGPRTPRRAAQAIALGAERKVEIFYRRDSCFCIVDVVVKEDG 120
 DB 61 QRPCTGFLGPRTPRRAAQAIALGAERKVEIFYRRDSCFCIVDVVVKEDG 120
 QY 61 QRPCTGFLGPRTPRRAAQAIALGAERKVEIFYRRDSCFCIVDVVVKEDG 120
 DB 61 QRPCTGFLGPRTPRRAAQAIALGAERKVEIFYRRDSCFCIVDVVVKEDG 120
 QY 121 AVIMEILNFVMEKMDVGPADNHRGPTSMABGRKTERLKLPAALLATARESSV 180
 DB 121 AVIMEILNFVMEKMDVGPADNHRGPTSMABGRKTERLKLPAALLATARESSV 180
 QY 121 AVIMEILNFVMEKMDVGPADNHRGPTSMABGRKTERLKLPAALLATARESSV 180
 DB 121 AVIMEILNFVMEKMDVGPADNHRGPTSMABGRKTERLKLPAALLATARESSV 180
 QY 181 RSGGAGAGAPGAVVDVLTTPAAPSSSLALDEVYTMNDHNVAGLGAERRALVGGSP 240
 DB 181 RSGGAGAGAPGAVVDVLTTPAAPSSSLALDEVYTMNDHNVAGLGAERRALVGGSP 240
 QY 155 -GSGGSHSRP-----SGGSP-----GGRG 173
 DB 155 -GSGGSHSRP-----SGGSP-----GGRG 173
 QY 241 PRSAPGLPSPRAHSLNPDAAGSSCSLARTSRSCASVRRASSADDIEMRAGVLRPP 300
 DB 241 PRSAPGLPSPRAHSLNPDAAGSSCSLARTSRSCASVRRASSADDIEMRAGVLRPP 300
 QY 174 ----- 173
 DB 174 ----- 173
 QY 301 RHAAGAMHPLRSLGSLNSTDSDLVRYRTISKIPQITLNFVDLKGDFLSPDSREITIA 360
 DB 301 RHAAGAMHPLRSLGSLNSTDSDLVRYRTISKIPQITLNFVDLKGDFLSPDSREITIA 360
 QY 174 -----KYRTSQIQPTLNFVEFULEKHSRSGSTTEIEIIA 208
 DB 174 -----KYRTSQIQPTLNFVEFULEKHSRSGSTTEIEIIA 208
 QY 361 P-KIKERTHWTEKVTQVLSIGADVLEPKYKQARIRHRTILHSPKAWMDLILLYI 419
 DB 361 P-KIKERTHWTEKVTQVLSIGADVLEPKYKQARIRHRTILHSPKAWMDLILLYI 419
 QY 209 PKVYERTQNTKVTQVLSIGADVLEPKYKQARIRHRTILHSPKAWMDLILLYI 268
 DB 209 PKVYERTQNTKVTQVLSIGADVLEPKYKQARIRHRTILHSPKAWMDLILLYI 268
 QY 420 YTAVFETYSAAFLKTEEGPATEGACOPLVVDLIDIMFIVILINFRTYYNAN 479
 DB 420 YTAVFETYSAAFLKTEEGPATEGACOPLVVDLIDIMFIVILINFRTYYNAN 479
 QY 269 YTAVFETYSAAFLKTEEGPATEGACOPLVVDLIDIMFIVILINFRTYYNAN 327
 DB 269 YTAVFETYSAAFLKTEEGPATEGACOPLVVDLIDIMFIVILINFRTYYNAN 327
 QY 480 EEVVSHBGRITAVHYFKGMFLIDVAAIPFDLLIFGSGSEE--DIGELKTARLLRYVA 536
 DB 480 EEVVSHBGRITAVHYFKGMFLIDVAAIPFDLLIFGSGSEE--DIGELKTARLLRYVA 536
 QY 328 DEVVSHBGRITAVHYFKGMFLIDVAAIPFDLLIFGSGDETTLLIGLKTARLLRYVA 387
 DB 328 DEVVSHBGRITAVHYFKGMFLIDVAAIPFDLLIFGSGDETTLLIGLKTARLLRYVA 387
 QY 537 RKLDRISEYGAAYFLMCTFPALIAHMLACIWAIGMNEOPHMSRIGMLHNLGDQIGP 596
 DB 537 RKLDRISEYGAAYFLMCTFPALIAHMLACIWAIGMNEOPHMSRIGMLHNLGDQIGP 596
 QY 388 RKLDRISEYGAAYFLMCTFPALIAHMLACIWAIGMNEOPHMSRIGMLHNLGDQIGP 447
 DB 388 RKLDRISEYGAAYFLMCTFPALIAHMLACIWAIGMNEOPHMSRIGMLHNLGDQIGP 447
 QY 597 YNSSG-LGSPRIKRYTALYFPPSSLSVSGFVNSPNTSEKIFESICVMLIGSLMYASI 655
 DB 597 YNSSG-LGSPRIKRYTALYFPPSSLSVSGFVNSPNTSEKIFESICVMLIGSLMYASI 655
 QY 448 YNSGDPASGSPVQKRYTALYFPPSSLSVSGFVNSPNTSEKIFESICVMLIGSLMYASI 507
 DB 448 YNSGDPASGSPVQKRYTALYFPPSSLSVSGFVNSPNTSEKIFESICVMLIGSLMYASI 507
 QY 656 FGNVSAITIORLYSGTARYHTQMLRVREFIRHQIPNPLRQLEIEYFOHMSYTGIDMNA 715
 DB 656 FGNVSAITIORLYSGTARYHTQMLRVREFIRHQIPNPLRQLEIEYFOHMSYTGIDMNA 715
 QY 508 FGNVSAITIORLYSGTARYHTQMLRVREFIRHQIPNPLRQLEIEYFOHMSYTGIDMNA 567
 DB 508 FGNVSAITIORLYSGTARYHTQMLRVREFIRHQIPNPLRQLEIEYFOHMSYTGIDMNA 567
 QY 716 VLKGFPECLQADICLHLNRSLLQCKPRGATKCLRLAKKFTTNAPEGDTLVHAGDL 775
 DB 716 VLKGFPECLQADICLHLNRSLLQCKPRGATKCLRLAKKFTTNAPEGDTLVHAGDL 775
 QY 568 VLKGFPECLQADICLHLNRSLLQCKPRGATKCLRLAKKFTTNAPEGDTLVHAGDL 627
 DB 568 VLKGFPECLQADICLHLNRSLLQCKPRGATKCLRLAKKFTTNAPEGDTLVHAGDL 627
 QY 776 LTAIFYTSGSIELLRGVVVAALLGKNDIEEPNLIVARPKSGNDVVALYCOLHKIR 835
 DB 776 LTAIFYTSGSIELLRGVVVAALLGKNDIEEPNLIVARPKSGNDVVALYCOLHKIR 835
 QY 628 LSTLYFTSRGSIELLRDVVVAALLGKNDIEEPNLIVARPKSGNDVVALYCOLHKIR 687
 DB 628 LSTLYFTSRGSIELLRDVVVAALLGKNDIEEPNLIVARPKSGNDVVALYCOLHKIR 687
 QY 836 DLLLEVLDVMEYFSDHFWSSLEITFNLDYV---MTP-GSPGTELEGFSRQKRKLS 890
 DB 836 DLLLEVLDVMEYFSDHFWSSLEITFNLDYV---MTP-GSPGTELEGFSRQKRKLS 890
 QY 688 ADLLEVDVMDYAFADTFMNKLEVFYNLRDADAGGLQSTPRQAPGHQDPQGEF----- 738
 DB 688 ADLLEVDVMDYAFADTFMNKLEVFYNLRDADAGGLQSTPRQAPGHQDPQGEF----- 738
 QY 891 FRRRTDKDTEQPGVYSALGPGACAGPSSRGRPGGPMGESPSSGPPSPESSEDEBPGRSS 950
 DB 891 FRRRTDKDTEQPGVYSALGPGACAGPSSRGRPGGPMGESPSSGPPSPESSEDEBPGRSS 950
 QY 739 -----LNDQSGAAPS----- 749
 DB 739 -----LNDQSGAAPS----- 749

DT 28-FEB-2003 (Rel. 41, last sequence update)
 DT 28-FEB-2003 (Rel. 41, last annotation update)
 DE Potassium voltage-gated channel subfamily H member 4 (Ether-a-go-go-
 DE like potassium channel 1) (Elk channel 1) (Elk1) (Brain-specific eag-
 DE like channel 2) (BCC2).
 GN KCNH4.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RA MEDLINE=9386988; PubMed=10455180;
 RA Miyake A., Mochizuki S., Yokoi H., Kohda M., Furuchi K.;
 RT "New ether-a-go-go K⁺ channel family members localized in human
 RT telencephalon.";
 RL J. Biol. Chem. 274:25018-25025(1999).
 CC -1- FUNCTION: Pore-forming (alpha) subunit of voltage-gated potassium
 CC channel. Elicits an outward current, but shows no inactivation.
 CC Channel properties may be modulated by cAMP and subunit assembly.
 CC -1- SUBUNIT: The potassium channel is probably composed of a homo- or
 CC heterotetrameric complex of pore-forming alpha subunits that can
 CC associate with modulating beta subunits.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- TISSUE SPECIFICITY: Detected only in brain, in particular in the
 CC telencephalon. Detected in putamen and caudate nucleus, and at
 CC lower levels in cerebral cortex, occipital and hippocampus.
 CC -1- DOMAIN: The segment S4 is probably the voltage-sensor and is
 CC characterized by a series of positively charged amino acids at
 CC every third position.
 CC -1- SIMILARITY: Belongs to the potassium channel family. H (Eag)
 CC subfamily.
 CC -1- SIMILARITY: Contains 1 cyclic nucleotide-binding domain.
 CC -1- SIMILARITY: Contains 1 PAS (PER-ARNT-SIM) dimerization domain.
 CC -1- SIMILARITY: Contains 1 PAS-associated C-terminal (PAC) domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: AB022698; BA83592.1; -
 DR Genew; HENC:6253; KCNH4.
 DR MIM; 604528; -
 DR GO: GO:0008076; C:voltage-gated potassium channel complex; TAS.
 DR GO: GO:0005249; F:voltage-gated potassium channel activity; TAS.
 DR GO: GO:0006813; F:potassium ion transport; TAS.
 DR InterPro: IPR000595; CNMP-binding.
 DR InterPro: IPR005821; Ion_trans.
 DR InterPro: IPR001622; K+channel_pore.
 DR InterPro: IPR005820; M+channel_mlg.
 DR InterPro: IPR001610; PAC.
 DR InterPro: IPR000700; PAS-assoc_C.
 DR InterPro: IPR000014; PAS_domain.
 DR Pfam: PF00027; CNMP-binding; 1.
 DR Pfam: PF00520; Ion_trans; 1.
 DR Pfam: PF00785; PAC; 1.
 DR SMART; SM00100; CNMP; 1.
 DR SMART; SM00086; PAC; 1.
 DR SMART; SM00091; PAS; 1.
 DR TIGRFAMs; TIGR00229; sensory_box; 1.
 DR PROSITE; PS50042; CNMP_BINDING_3; 1.
 DR PROSITE; PS50112; PAC; 1.
 DR PROSITE; PS50113; PAC; 1.
 KM Transport; Ion transport; Ionic channel; Voltage-gated channel;
 KM Potassium channel; Potassium; Potassium transport; Transmembrane;
 KM Glycoprotein; Multigene family.
 FT DOMAIN 1 228 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 229 249 SEGMENT S1 (POTENTIAL).

FT TRANSMEM 260 280 SEGMENT S2 (POTENTIAL).
 FT DOMAIN 281 302 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 303 323 SEGMENT S3 (POTENTIAL).
 FT TRANSMEM 333 353 SEGMENT S4 (POTENTIAL).
 FT DOMAIN 354 361 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 362 382 SEGMENT S5 (POTENTIAL).
 FT DOMAIN 428 448 SEGMENT H5 (PORE-FORMING) (POTENTIAL).
 FT TRANSMEM 483 503 SEGMENT S6 (POTENTIAL).
 FT DOMAIN 504 1017 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 14 90 PAS.
 FT DOMAIN 93 145 CNMP.
 FT NP_BIND 556 671 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 326 326 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 414 414 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 473 473 N-LINKED (GLCNAC. .) (POTENTIAL).
 SQ SEQUENCE 1017 AA; 111692 MW; 48480DECT75810A38 CRC64;
 Query Match 26.9%; Score 1636; DB 1; Length 1017;
 Best Local Similarity 34.9%; Pred. No. 1.8e-77;
 Matches 424; Conservative 153; Mismatches 319; Indels 318; Gaps 41;
 QY 1 MPVRGVAQONTFLDTITIRKFEQSRRFIITANAR-VCNCAVITCNGFCGLGYSRAEV 59
 D 1 MPVKKGLAQNFTFLDIATRFDTGHSNFTLANAQTRGPPIYCSDFCELTGGRTEV 60
 QY 60RRCCTDPLHGPRTORAAQIOALLGAEREVEITAFYRKDSGFCGLDVYPRKND 119
 D 61MOKTCSRFLGCPPTSEPALORLKLAEHDEHAEICFRKDSATWCLIDMPPIKNE 120
 QY 120GAVIMFLNFEVNEKMDVSPADHTNHRGPTSWLAPRAKTRKLPALALTARESS 179
 D 121GEVVLFLFSFRDIQS--GSPG-----LDP----- 143
 QY 180VRSGGAGGACAPGAVVVDVLTTPAIPSESLALDEVTAAMNHVAGLCPAERRALVPGS 239
 D 144--GGGRDS-----NHNESLG----- 157
 QY 240PPRSPAGQLSPRAHSLNPDASGSSCLARTRSRSCASVARRASADDIEMRAGVLP 299
 D 158-RRGATWKRFS-----ARRRST-----VL--- 176
 QY 300PRHASTGAMPLRSGSLNSTSDLVRYRTISKIPQITLNFVDLGDGDFLTPASDREII 359
 D 177--HRLTG--HGRRG-----GQGMANNVF 198
 QY 360APKIKERTHNVEKTVGLSLGADVLPDEYKLOAPRIHRTILHSPKRAVMDLILVI 419
 D 199EPK-----PSVEPKYVASVSGSRCLLHYSVSKAIDGILLATF 238
 QY 420YTANFEPYSAFLKETNEEGPAPTEGCGYACQPLAVVDLIDIMVILINERTTYNAN 479
 D 239YVAATVPYVWCF--SGDDDPITSS--RHLVSDIAVEMLILIDILIFRTTYVSQS 290
 QY 480EEVYSHPRGIAVHYFKGMFLIDWVAIPFDLL-IFGSGSEELIGLKTARLLRLVYARK 538
 D 291GVYISAPRSIGLHAWFLFDLIALPFDLITFNTIVSVLHLLTVRLRLRLRLOK 350
 QY 539LDRTSEYGAVALFLKCTFALIAHWLACIWAIGNMQPHMDS---RIGWLHNGDQIGK 595
 D 351LERYSCSAVALTLTLMGVFALLAHMMAICWIVIGIRREMANDPLMDIGWHEHGKRLV 410
 QY 596PYNSGGGSPSIRKQYVATLYEFPSSILTSVGFQVNSVNTSEKIFESICVMLIGSLMYASI 655
 D 411PYVNGSVGSPSRSAVTAALFTLSSLTSGVGVNCANTDAEKTFSTCTMLIGALMAV 470
 QY 656FGVNSAIIORLYSGTARVHTOMLRVREFIREHQIPNLRJORELEFYOHANSYTGIDMNA 715
 D 471FGVNTAIIORLYSGTARVHTOMLRVREFIREHQIPNLRJORELEFYOHANSYTGIDMNA 715
 QY 716VLKGFPECLADICLIHNSRLDHCPRRGATKCCALALAKRTTTHAPREDTLVHAGDL 775
 D 531LIRDFPELRADIAHMLNREITLQ-LPLFGAASRCLRALSLHKTSCARCAGEVLLRRGDA 589


```

QY 776 LTALYFISRGSEILLRGDVAVALILKNDIFG-----EPLNLYARPG---KSGNDVRA 824
DB 590 LQAHYVCGSLDEVLRLDMNVLAIIICKGDLIGADIDEPGQEP-GLGADPNFVLTQTSADVKA 648
QY 825 LTYCGLHRIHRDLDLEVDWPEFSDHFWSSL-ETTFNLRTNMIPGSPGSTELEGGS 882
DB 649 LTYCGLQQLSSNGLALEVLRLYREYGAARRAGLPDITFLKRO-----GSDTS-----GIS 698
QY 883 R-QRRRKLSFRRR---TDKTDQEPGEVSALGPRGAGAPSSR-----GRPGG-- 925
DB 699 RFSRGPRLSQPSESLSSSDKTLPSITEAESGAPGGPRRRPRLPLNLSPARPGSL 758
QY 926 -----PMG-----ESPSSGGS-SP-----ESEDDEPGSSS-----PLALVPPSSSR 962
DB 759 VSLGEEPLPPEFSAVSSPSLSPSLPALAGOGHSASPHPCPSAAMKPPOLLIP---PL 815
QY 963 PGEPGPGGP-EMEDCEKSPDTCNPLSGAFSGVSNFSPWSDRGRQOELPRCPAPPS 1021
DB 816 GTFGPPDLSPRIVDIEDSGTAEPFRFS-----RRPELPRPSQAP- 859
QY 1022 LNLIPSSPGRRPRGDVESRLDALQRLNRL-----TRLSDMATVLOLLQRCM 1071
DB 860 -----PTGTRPSEPLASAEVEKVCRLNOEISRLNOEVSQLSRELRIHMLGLQAR- 911
QY 1072 TLVPRAYSAVTPPGGPTSTPLVSPPTITLDSLSVSPFACCELPAPAPLPDGG 1131
DB 912 -LGPPGHPASGAWTPDP-----PCQPLRPPLGSL-----PCASRP--PSL-ODT 951
QY 1132 PTRRLSLPGOLGAL 1145
DB 952 TLAEVHCPASVGTM 965

```

RESULT 13

```

ID KCH4_RAT STANDARD: PRT: 1017 AA.
AC 09RT9: 089048:
DT 28-FEB-2003 (rel. 41, Created)
DT 28-FEB-2003 (rel. 41, Last sequence update)
DT 28-FEB-2003 (rel. 41, Last annotation update)
DE Potassium voltage-gated channel subfamily H member 4 (Ether-a-go-go-
DE like potassium channel 1) (ELK channel 1) (ELK1) (Brain-specific eag-
DE like channel 2) (BEC2).
GN KCN4 OR ELK1 OR ELK3.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A., AND TISSUE SPECIFICITY.
RC TISSUE=Brain Cortex;
RA MEDLINE=99043952; PubMed=9824707;
RA Engelard B., Neu A., Ludwig J., Roeper J., Pongs O.;
RT "Cloning and functional expression of rat ether-a-go-go-like K+
RT channel genes.";
RL J. Physiol. (Lond) 513:647-654(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RA MEDLINE=99386988; PubMed=10455180;
RA Miyake A., Mochizuki S., Yokoi H., Kohda M., Furuchi K.;
RT "New ether-a-go-go K+ channel family members localized in human
RT telencephalon.";
RL J. Biol. Chem. 274:25018-25025(1999).
RN [3]
RP TISSUE SPECIFICITY.
RA MEDLINE=21319165; PubMed=11425889;
RA Saganich M.J., Machado E., Rudy B.;
RT "Differential expression of genes encoding subthreshold-operating
RT voltage-gated K+ channels in brain.";
RL J. Neurosci. 21:4609-4624(2001).
CC -I- FUNCTION: Pore-forming (alpha) subunit of voltage-gated potassium
CC channel. Elicits an outward current, but shows no inactivation.

```

```

CC Channel properties may be modulated by cAMP and subunit assembly.
CC -I- SUBUNIT: The potassium channel is probably composed of a homo- or
CC heterotrimeric complex of pore-forming alpha subunits that can
CC associate with modulating beta subunits.
CC -I- SUBCELLULAR LOCATION: Integral membrane protein.
CC -I- TISSUE SPECIFICITY: Highly expressed in adult testis, and in adult
CC and embryonic brain. In adult brain found in piriform cortex,
CC olfactory tubercle, cerebral cortex, hippocampus pyramidal cells
CC and dentate gyrus and basal ganglia of caudate/putamen and
CC accumbens nucleus. Detected at intermediate levels in lung, spinal
CC cord, and pituitary.
CC -I- DEVELOPMENTAL STAGE: Expressed at day E18 in embryonic brain.
CC -I- DOMAIN: The segment S4 is probably the voltage-sensor and is
CC characterized by a series of positively charged amino acids at
CC every third position.
CC -I- SIMILARITY: Belongs to the potassium channel family. H (dag)
CC subfamily.
CC -I- SIMILARITY: Contains 1 cyclic nucleotide-binding domain.
CC -I- SIMILARITY: Contains 1 PAS (PER-ARNT-SIM) dimerization domain.
CC -I- SIMILARITY: Contains 1 PAS-associated C-terminal (PAC) domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AJ007628; CA07587.1; -.
DR EMBL: AB022699; BA083593.1; -.
DR PIR: T31354; T31354.
DR InterPro: IPR000595; CNMP_binding.
DR InterPro: IPR005821; Ion_trans.
DR InterPro: IPR001622; K+channel_pore.
DR InterPro: IPR005820; M+channel_nlg.
DR InterPro: IPR001610; PAC.
DR InterPro: IPR000700; PAS-assoc_C.
DR InterPro: IPR000014; PAS_domain.
DR Pfam: PF00027; CNMP_binding; 1.
DR Pfam: PF00520; Ion_trans; 1.
DR Pfam: PF00785; PAC; 1.
DR SMART: SM00100; CNMP; 1.
DR SMART: SM00086; PAC; 1.
DR SMART: SM00091; PAS; 1.
DR TIGRFAMs: TIGR00229; sensory_box; 1.
DR PROSITE: PS00042; CNMP_BINDING_3; 1.
DR PROSITE: PS01112; PAS; 1.
DR PROSITE: PS01113; PAC; 1.
KW Transport; Ion transport; Ionic channel; Voltage-gated channel;
KW Potassium channel; Potassium transport; Transmembrane;
KW Glycoprotein; Multigene family.
FT DOMAIN 1 232
FT TRANSMEM 263 283
FT TRANSMEM 263 283
FT TRANSMEM 284 305
FT TRANSMEM 306 326
FT TRANSMEM 335 355
FT DOMAIN 356 364
FT TRANSMEM 365 385
FT DOMAIN 428 448
FT TRANSMEM 455 475
FT DOMAIN 476 1017
FT DOMAIN 14 90
FT DOMAIN 93 145
FT NP_BIND 557 672
FT CARBOHYD 415 415
FT CONFLICT 825 825
SQ SEQUENCE 1017 AA; 111403 MW; BFD80F1B35437C9F CRC64;

```

Query Match 26.6%; Score 1616; DB 1; Length 1017;
 Best Local Similarity 33.6%; Pred. No. 1,9e-76;
 Matches 417; Conservative 172; Mismatches 341; Indels 312; Gaps 36;

DR SMART; SM00086; PAC; 1.
 DR TIGRFS; TIGR00229; sensory_box; 1.
 DR PROSITE; PSS0042; CNMP_BINDING_3; 1.
 DR PROSITE; PSS0113; PAC; 1.
 DR PROSITE; PSS0112; PAS; FALSE NEG.
 KW Transport: Ion transport; Ionic channel; Voltage-gated channel;
 KW Potassium channel; Potassium transport; Transmembrane;
 KW Glycoprotein; Multigene family.
 FT DOMAIN 1 225
 FT TRANSMEM 226 246
 FT TRANSMEM 256 276
 FT TRANSMEM 277 298
 FT TRANSMEM 299 319
 FT TRANSMEM 328 348
 FT TRANSMEM 349 357
 FT TRANSMEM 358 378
 FT TRANSMEM 420 440
 FT TRANSMEM 449 469
 FT TRANSMEM 470 1107
 FT DOMAIN 18 90
 FT DOMAIN 93 145
 FT DOMAIN 951 1064
 FT DOMAIN 711 723
 FT NP_BIND 551 668
 FT CARBOHD 320 320
 FT CARBOHD 409 409
 SQ SEQUENCE 1107 AA; 123832 MW; 4DC93EF85B674905 CRC64;

Query Match 25.8%; Score 1568; DB 1; Length 1107;
 Best Local Similarity 33.8%; Pred. No. 6.5e-74;
 Matches 384; Conservative 163; Mismatches 298; Indels 292; Gaps 27;

QY 1 MPVRGHVAPONTFLDITIRKEGOSRKFIANARV-ENCAYICNDGFCCLGYSAEV 59
 DB 1 MPVMGLAPQWTFEDTATRPDGTSHSNFIANAQVANGFPIYVCGSCGELAGARREV 60
 QY 60 MQRCTCPFLHGRPRQRAAOIAOALLGAERKEVIAFYKDGSCFLCVDVYVKNED 119
 DB 61 MOKSCSCFLFGEVTEQMLQIEKSELEKTEFKELMEFYKNGSPFCLDIVIKNEK 120
 QY 120 GAVIMEINFEVYMEKDWGSPAHDTNRGPPTSMLAGPRAKTFPLKLPALLALARESS 179
 DB 121 GDVVFLELSF-----KDT-----DTK----- 137
 QY 180 VRSGAGAGAGAVVVDVLTLPAPSSSESLADEVTAMDNHVAGIGPAEERRALVPGSS 239
 DB 138 -----VKITP-----EDKKEDKVK----- 151
 QY 240 PPRSAFGDLPBRAHSLMPDASGSSCSIAKTRSRSCASVRRASSADIDEAMRAGVLP 299
 DB 152 -----GRSRAGTHFDSARRRSRA----- 169
 QY 300 PRHASTGAMHPLRSLGSLNSTSDSLVRYRTISKIQITLFPDLGDPFLASPTSDRELI 359
 DB 170 -----VLHISHGLDRREKNIKLIKINNVPVDPK----- 197
 QY 360 APKIKERTHNTEKVTQVLSGADVLPEYKLOAPRIHMTKILHSPFAVMDLILLYVT 419
 DB 198 -----AFPEYKVSDAKSKFILLHFSTFAGMDWILLATF 233
 QY 420 YTAFTPTPSAFLKETEGRPAATCGACOPLAVDLVIDIMFVLDILINERTTYVAN 479
 DB 234 YVAATVPYVNCFIGNDLSTRTST-----TVSDAVELLEFIIDILINERTTYVKS 284
 QY 480 EEVYSHPRIVAVHFKEGFLDMVAALPFDLL-IFGSSSEELIGLKTARILRIYVAR 538
 DB 285 GOVTFEARSICIRHYVTWFTIDLAALPFDLLVAVNTVSVLHLKTVRLRLRLLOK 344
 QY 539 LDRYSEGAVALFLIMCTFALIAHLCIWAIGMEOPH--MDSRIGWLHNLDOIQK 595
 DB 345 LDRYSQHSHTYVLTILMSFALLAHMACIWIYIGMEEDSLKWEYGLHETLOKRLDS 404
 QY 596 P-YNSGLGSPISDKYVTAIFYTFSSILTSVGCVNSPNTNSEKIFISICVMLIGSLMYAS 654

DB 405 PYGNNLTGCGPISRASYIALALFTLLSLTSVGFVGNVSANTDEKIFSLITMIGALMHAL 464
 QY 655 IFGNYSATIORLGYSGTARHYTOMLNRREYTRHOJPNPLROBLEYFQOAMSYTNGIDNN 714
 DB 465 VEGNVTAILIQRYSTRKSLYHRTKDKIDIRVHNHLPDOLKQMLEYFOTTMYSVNNGIDSN 524
 QY 715 AVLKGFPECLQADICLHLNRSLLQHCPRGATKCLRALAMKFTTHAPPGDITLVHACD 774
 DB 525 ELKDPFDELRSDITMHLNKETILQ-LSLEFCASRGCLRSLSHKITSFCAPEGVLLROGD 583
 QY 775 LITALEFISRGSEIELRGDVVAIIIGKNDIEPEPLNLYARPKSNGDVALTYCDLAKH 834
 DB 584 ALQATFVCGSGMEVLCOSMVALIIGKGLIGANLSIKQVYKTNADVALLTYCDLQCI 643
 QY 835 RDOLLEVDLMYEFSDHFESSL-ETTFELRD-----TNMIPGSGTELEGGSRQR 885
 DB 644 LKGLFEVLDLIFYEYAKHFVEDIQHDLITLYLRGCHESDVYLSLNSKMSVSOSEPKNGNIN 703
 QY 886 KRKLSFRRTDKDQOPG--EVSALGPRAGAGPSSRGPRG-----PWGE 929
 DB 704 KRLPSTIVE--DEEEDEEGEEAVSLSP-ICTRGSSSRKKVGSNKAYLGLSKQLASGT 760
 QY 930 SPSSGP-----SPESSEDEGPRGSS-----PLRLVPESSPRPPEPGEPLMECC 977
 DB 761 VFPHSPIRVSRSNSPKTQKQELDPNNHKKREKNLKLQISTLNACGPPDISPRIVDIGEDG 820
 QY 978 EKSPDTCNPLSGAFSGVSNIFSFMGDSRGQOELPRCAPPSLLNIPLSGPRGRPD 1037
 DB 821 NSSEE-----SOTFDF-GSERIR-----SEPRISP---PLGD 848
 QY 1038 VE-----SRDLALQOLNRL-----TRLSADMATVLOLLOMTLVPPA 1077
 DB 849 PEIGAALVFIKAEETKQKQINKLNSEVYTLQEVSQLGKQMRNVIRLENVLSPPQPS 905

RESULT 15
 KCH8_RAT STANDARD; PRT; 1102 AA.
 AC 09QWS8; 088877;
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Potassium voltage-gated channel subfamily H member 8 (Ether-a-go-go-like potassium channel 3) (ELK channel 3).
 GN KCNH8 OR ELK3 OR ELK1.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=98382545; PubMed=9714851;
 RA Shi W., Wang H.-S., Pan Z., Wymore R.S., Cohen I.S., McKinnon D., Dixon J.E.;
 RT "Cloning of a mammalian elk potassium channel gene and ERG mRNA distribution in rat sympathetic ganglia."
 RN [2]
 RP J. Physiol. (Lond) 511:675-682(1998).
 RP SEQUENCE OF 9-379 FROM N.A.
 RC TISSUE=Brain cortex;
 RX MEDLINE=99043952; PubMed=9824707;
 RA England B., Neu A., Ludwig J., Reeper J., Pongs O.;
 RT "Cloning and functional expression of rat ether-a-go-go-like K+ channel genes."
 RN [3]
 RP J. Physiol. (Lond) 513:647-654(1998).
 RP TISSUE SPECIFICITY.
 RX MEDLINE=20183472; PubMed=10718922;
 RA Wulfsen I., Hauber H.-P., Schlemann D., Bauer C.K., Schwarz J.R.;
 RT "Expression of mRNA for voltage-dependent and inward-rectifying K channels in GH3/B6 cells and rat pituitary."
 RL J. Neuroendocrinol. 12:263-272(2000).

RN [4]
 RP TISSUE SPECIFICITY.
 RX MEDLINE-21319165; PubMed-11425889;
 RA Saganich M.J., Machado E., Rudy B.;
 RT "Differential expression of genes encoding subthreshold-operating
 RT voltage-gated K⁺ channels in brain."
 RL J. Neurosci. 21:4609-4624(2001).
 CC -1- FUNCTION: Pore-forming (alpha) subunit of voltage-gated potassium
 CC channel. Elicits a slowly activating, outward rectifying current.
 CC Channel properties may be modulated by cAMP and subunit assembly.
 CC -1- SUBUNIT: The potassium channel is probably composed of a homo-
 CC tetramer with modulating beta subunits.
 CC -1- TISSUE SPECIFICITY: Detected in superior cervical, mesenteric and
 CC coeliac ganglia. Expressed in brain (piriform cortex, olfactory
 CC tubercle, cerebral cortex, hippocampus pyramidal cells and
 CC dentate gyrus and basal ganglia of caudate/putamen and accumbens
 CC (nucleus). Expressed in pituitary.
 CC -1- DOMAIN: The segment S4 is probably the voltage-sensor and is
 CC characterized by a series of positively charged amino acids at
 CC every third position.
 CC -1- SIMILARITY: Belongs to the potassium channel family. H (Bag)
 CC subfamily.
 CC -1- SIMILARITY: Contains 1 cyclic nucleotide-binding domain.
 CC -1- SIMILARITY: Contains 1 PAS (PER-ARNT-SIM) dimerization domain.
 CC -1- SIMILARITY: Contains 1 PAS-associated C-terminal (PAC) domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AF061957; AAC61520.1; -
 DR EMBL; AJ007632; CA07591.1; -
 DR PIR; T17367; T17367.
 DR InterPro: IPR000595; cNMP_binding.
 DR InterPro: IPR003967; Erg_channel.
 DR InterPro: IPR005821; Ion_channel.
 DR InterPro: IPR003280; K+channel_2pore.
 DR InterPro: IPR001622; K+channel_pore.
 DR InterPro: IPR005820; M+channel_nlg.
 DR InterPro: IPR001610; PAC.
 DR InterPro: IPR000700; PAS_assoc_C.
 DR InterPro: IPR000014; PAS_domain.
 DR Pfam; PF00027; cNMP_binding; 1.
 DR SMART; SM00100; cNMP; 1.
 DR SMART; SM00086; PAC; 1.
 DR TIGRfam; TIGR00229; sensory_box; 1.
 DR PROSITE; PS0042; cNMP_BINDING_3; 1.
 DR PROSITE; PS0113; PAC; 1.
 DR PROSITE; PS0112; PAS; FALSE_NEG.
 KW Transport: Ion transport; Potassium channel; Voltage-gated channel;
 KW Potassium channel; Potassium; Potassium transport; Transmembrane;
 KW Glycoprotein; Multigene family.
 KM
 FT DOMAIN 1 225
 FT TRANSMEM 226 246
 FT TRANSMEM 256 276
 FT DOMAIN 277 298
 FT TRANSMEM 299 319
 FT TRANSMEM 328 348
 FT TRANSMEM 349 353
 FT TRANSMEM 354 374
 FT DOMAIN 420 440
 FT TRANSMEM 449 469
 FT DOMAIN 470 1102
 FT DOMAIN 18 90
 FT DOMAIN 93 145
 FT DOMAIN 711 722
 FT POLY-GLU.

FT NP_BIND 551 668 CNMP.
 FT CARBOHYD 320 320 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 409 409 E -> L (IN REF. 2).
 FT CONFLICT 71 71 F -> N (IN REF. 2).
 FT CONFLICT 187 187 K -> N (IN REF. 2).
 FT CONFLICT 296 296 I -> T (IN REF. 2).
 FT CONFLICT 370 370 M -> I (IN REF. 2).
 SQ SEQUENCE 1102 AA; 123230 MW; A135CC36E2E7F1A3 CRC64;
 Query Match 25.5%; Score 1551; DB 1; Length 1102;
 Best Local Similarity 32.6%; Pred. No. 4.9e-73;
 Matches 383; Conservative 169; Mismatches 283; Indels 340; Gaps 30;
 QY 1 MPVRGVAPQNFELDTIIRKFEQSGSKFTIINARV-ENCVAIYCNDFCELCYSRAEV 59
 DB 1 MPVKGILAFQNFELDTIIRFEGTSHNFTLIANAQVAKGPPIYCSGFEELAGFATEV 60
 QY 60 MRRPCTDFLHGRPTORRAAQAQALGAERKEVETAFYRKDSCFLVDVVPVKNED 119
 DB 61 MRSCKCFLEGVETNQMLQIEKSLDEKVEFKEGEMFYKKNAGPFWCLDIPIKNEK 120
 QY 120 GAVIMPLINEVYEMKDWGSPADHTNHRPPTISLAPGAKTFRLKPLALITARESS 179
 DB 121 GDVVLPLASF-----KDIPTKY----- 138
 QY 180 VRSGAGAGACAPGAVVVDVLDTPAAPSSSLADEVTAMDNVHAGLGPAPERALVPGS 239
 DB 139 -----KITSEK-----KEDR----- 149
 QY 240 PRPAPQQLPSPRAHSLNPDSGSSCLARTRSRSCASVRRASADDIEMRAGVLP 299
 DB 150 -----AKGR-----SRAGSHFDSARRSR----- 168
 QY 300 PRASIGAMPLSLGSLNSTSDSLVYRTISKIPQTLNFDVLDGEPFLASPSDRETI 359
 DB 169 -----AVLYHISGHLDREKKRLKINNVPFDKP----- 197
 QY 360 APRKERTHNTEKVTQVLSGADVLDPEYKIQAPRIHRTMILHSPKAWMDLILLYI 419
 DB 198 -----APPEYKVSDAKKKFKFLHSTFKACMDLILATF 233
 QY 420 YTAFTPYSAFLFKETEGEPATEGCGYACQPLAVVDLIVDIMEYDILINFRTYVAN 479
 DB 234 YVAVTPVYVNCFIGNEDLSTTRST-----TVSDIAVEILFEIIDIILNFRITYVKS 284
 QY 480 EEVYSHRGRIAVYFKMFLIDWAAIPDOL-IFGSGSEELGLKTAALLRYVRARK 538
 DB 285 GGVIFEARSLCIHYVTWTFIIDIILALPFDLLAFNVTVVSVLHLKTVRLRLRLLOK 344
 QY 539 LDRYSEGAUVLEFLMCTFALIAHMLACIYVATGNMOPH---MDSRIQHLNLDGQIGK 595
 DB 345 LDRYSOHSTVYLLLSMFAILLHMMACIYVIGKMEREDNSLWKVEYGMHLHEIGKLES 404
 QY 596 P-YNSGSLGSPSIRKDYVTALEYTFSSLTSGVGNVSPNTNSEKIFSIYMLIGSLMYAS 654
 DB 405 PYTGNNLTIVLLLSMFAILLHMMACIYVIGKMEREDNSLWKVEYGMHLHEIGKLES 464
 QY 655 IFGNVSAIIRLISGARITYTQMLRYREFIRFQIIPPLQRLLEYFQHMWSTINGIDAN 714
 DB 465 VFGNVNVAIIIRMSRMSLYTRFKDLKDFRVNHLPOOLQRMLEYQTWSVNGIDSN 524
 QY 715 AVKGPPECIOADIILNSLSLOHCKPFGATKGCALAMKFKTHAPRGDTLVAGD 774
 DB 525 ELKDPDELRSDITHMLNKEIIO-LSLFECASGICRLSLNHRKTSFCAPGELLAQGD 583
 QY 775 LITALLYFISGSIIEILGVDVVAALIKNDIFGEPLNLVAPRGKSGNDVALTYCDLHKIH 834
 DB 584 ALQAIYFVCGSMELVLDKSDVLAALILGKDLIGANLSTIKDVITKNADVKALFYCDLCCII 643
 QY 835 RDQLLEVLDMPPEPSDHFWSL--EITFNLRD-----TNMIRPSPGSTELGEG 880
 DB 644 LKGLFEVLYGLYPRYAHKFVEDIQLHDTLYNLRGHESDVVISRLSNKSTVP---QAERKN 699

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 23, 2003, 11:05:09 ; Search time 88 Seconds
(without alignments)
3398.670 Million cell updates/sec

Title: US-10-000-151B-3
Perfect score: 6079
Sequence: 1 MPVRGHVAPQNTFLDITIR.....GOLGALTSQPLRHGSDPGS 1159

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

SPTREMBL_23:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mmc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_protent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirs:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5228.5	86.0	1058	4	Q81Z12
2	3241	53.3	1186	13	Q8UH78
3	2430	40.0	732	4	Q81V15
4	1941.5	31.9	791	5	Q18325
5	1865.5	30.7	855	5	Q02497
6	1530	25.2	515	11	Q8C782
7	1530	25.2	1097	11	Q8BX82
8	1516.5	24.9	522	11	Q8CC38
9	1485.5	24.4	1284	5	Q9V899
10	1485.5	23.9	1174	5	Q23974
11	1453.5	22.4	956	5	Q9YX26
12	1423	23.4	956	5	Q9YX7
13	1403.5	22.1	988	11	Q8C035
14	1099.5	18.1	395	11	Q8BY52
16	511.5	8.4	1327	5	Q9V702

17	503.5	8.3	678	5	Q96777	heliobis v
18	502.5	8.3	945	5	Q9Y1J9	09Y1J9 drosophila
19	459	7.6	113	11	Q8BRP8	08BRP8 mus musculus
20	440	7.2	767	5	Q76977	076977 strongyloce
21	439.5	7.2	900	5	Q97119	097119 limulus pol
22	434	7.1	1463	5	Q905E2	0905E2 drosophila
23	428.5	7.0	688	10	Q41461	041461 solanum tub
24	427	7.0	810	10	Q9XH40	09XH40 samanea sam
25	426	7.0	1696	5	Q9YXV8	09YXV8 drosophila
26	425	7.0	1324	5	Q8IR35	08IR35 drosophila
27	420.5	6.9	611	11	Q9QW7	09QW7 rattus norv
28	420.5	6.9	632	11	Q9ER32	09ER32 rattus norv
29	420.5	6.9	670	11	Q9ER33	09ER33 rattus norv
30	418.5	6.9	631	11	Q8CFV6	08CFV6 mus musculu
31	418.5	6.9	1218	5	Q9W201	09W201 drosophila
32	412.5	6.8	802	10	Q38898	038898 arabidopsis
33	412	6.8	787	10	Q3M0L6	03M0L6 arabidopsis
34	409	6.7	686	6	Q9N0H4	09N0H4 sus scrofa
35	397	6.5	883	10	Q38998	038998 arabidopsis
36	397	6.5	883	10	Q9LE66	09LE66 lycopersico
37	393	6.5	883	10	Q24382	024382 solanum tub
38	391.5	6.4	677	10	Q42426	042426 arabidopsis
39	389.5	6.4	677	10	Q3128	03128 arabidopsis
40	389.5	6.4	820	10	Q9FNY5	09FNY5 arabidopsis
41	387.5	6.4	830	10	Q9FY04	09FY04 populus tre
42	385	6.3	849	10	Q9SM12	09SM12 zea mays (m
43	385	6.3	887	10	Q04242	004242 zea mays (m
44	384.5	6.3	697	10	Q49732	049732 arabidopsis
45	384.5	6.3	708	10	Q9C5V9	09C5V9 arabidopsis

ALIGNMENTS

RESULT 1	ID	Q81Z12	PRELIMINARY;	PRT; 1058 AA.
AC	Q81Z12;			
DT	01-MAR-2003 (TREMBLrel. 23, Created)			
DT	01-MAR-2003 (TREMBLrel. 23, Last sequence update)			
DT	01-MAR-2003 (TREMBLrel. 23, Last annotation update)			
DE	Ether-a-go-go related potassium channel.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Colon;			
RA	Shoeb F., Malykhina A.P., Akbarali H.I.;			
RT	"HERG potassium channel from human colon."			
RL	Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.			
DR	EMBL; AY130462; RAN05415.1; -			
KW	Ionic channel.			
SQ	SEQUENCE 1058 AA; 116000 MW; 70F37F537AAFE084 CRC64;			
Query Match	Best Local Similarity	86.0%;	Score 5228.5;	DB 4; Length 1058;
Matches 1005;	Conservative	95.1%;	Pred. No. 0;	
	Mismatches	9;	Indels	5; Gaps 3;
QY	1	MPVRGHVAPQNTFLDITIRKFEKGRKFTIANARVENCAYIYCNDFGELGYSRAEYM 60		
DB	1	MPVRGHVAPQNTFLDITIRKFEKGRKFTIANARVENCAYIYCNDFGELGYSRAEYM 60		
QY	61	QRPCTDHLHGPRTORRAAQAIALGAEERKVEIAFYRKDGSCFLCLVDVVPVKNEDG 120		
DB	61	QRPCTDHLHGPRTORRAAQAIALGAEERKVEIAFYRKDGSCFLCLVDVVPVKNEDG 120		
QY	121	AIIMFVLNEVMEKDMGSSPAHDPNHRPSTSLAAGAKTFRKLPLALLTARESSV 180		
DB	121	AIIMFVLNEVMEKDMGSSPAHDPNHRPSTSLAAGAKTFRKLPLALLTARESSV 180		
QY	181	RSGAGAGAGAPGAVVVDVLTTPAAPSSESLADEVTAMDNHVAIGLPAREBRALVPG-- 238		

D	b	 RTGSMRSGAGVAVVDADLTTPAAPSSSLALDGVASANDNHVAGLPAEERRALLPGSQA	240
O	y		229 SPSPSAPOQLPSRAHSLNPDASSGCCLATRRSCASVARRASADDIEAMRGVLP	298
D	b		241 SPVASIRPGRPHSPRAQGLNPDAAGSCSCPARKTRHESCASVRASASADIEMARGALPP	300
O	y		299 PPRHASTGAMHPRLNSGLNSTSDSLVRYRTISKIPQTLTNFVDKGPFLASPDSDEI	358
D	b		301 PPRHASTGAMHPRLNSGLNSTSDSLVYKRITSKPQTLLNFVDLKGPFLASPDSDEI	360
O	y		359 IAPRIKERHTANVEKEVTQVLSLGADVLEPYKLQABRIHRWTLIHSPKCAVMDLILLY	418
D	b		361 IAPRIKERHTANVEKEVTQVLSLGADVLEPYKLQAPKIHRTWLTHSPKCAVMDLILLY	420
O	y		419 IYTVAFFPYSAAPFLIKETEEGPPEEGYACOPPLAVVDLIYDIMITVILLINFRTYYNA	478
D	b		421 IYTVAFFPYSAAPFLIKETEDGSQAPEGYACOPPLAVVDLIYDIMITVILLINFRTYYNA	480
O	y		479 NEEVYSHBGRIVAHYFKGMFLIDVAAIPFDLLIFGSGSEELIGLKTARLLRVARVK	538
D	b		481 NEEVYSHBGRIVAHYFKGMFLIDVAAIPFDLLIFGSGSEELIGLKTARLLRVARVK	540
O	y		539 LDRXSEYAVALFLMCTCFALIANTHACIWAIGMDEPPhMSDKRGMLHNIGDOIGKYNN	598
D	b		541 LDRXSEGAVALFLMCTCFALIANTHACIWAIGMDEPPhMSDKRGMLHNIGDOIGKYNN	600
O	y		599 SSGJGGPBIKDKXYVALFTFESSLTSVGFGVNSPPTNSEKIFPSICVMILGISMTASIIGN	658
D	b		601 SSGJGGPBIKDKXYVALFTFESSLTSVGFGVNSPPTNSEKIFPSICVMILGISMTASIIGN	660
O	y		659 VSAIIORLYSGTARYHTOMLVREFIREHQIPNPLROBLEEYFOHAMSYNGIDMNAVVK	718
D	b		661 VSAIIORLYSGTARYHTOMLVREFIREHQIPNPLROBLEEYFOHAMSYNGIDMNAVVK	720
O	y		719 GFPECLQADICHLNRSLLJOHCXPKPRAGTKGCLRALKMKFKTTNARPDDTLVHAGDLTA	778
D	b		721 GFPECLQADICHLNRSLLJOHCXPKPRAGTKGCLRALKMKFKTTNARPDDTLVHAGDLTA	780
O	y		779 LYFSIRGSIETILRGDVAAVLIGKNDIFEGBPLNLVARPKSGNDVATLYCDLHKIHRDDL	838
D	b		781 LYFSIRGSIETILRGDVAAVLIGKNDIFEGBPLNLVARPKSGNDVATLYCDLHKIHRDDL	840
O	y		839 LEVIDMWPEFSDHFMWSLETFENLRDTMMIRIGSPSTLEDGFSSQRKRKISFRRTXTKD	898
D	b		841 LEVIDMWPEFSDHFMWSLETFENLRDTMMIRIGSPSALSLEGFPNQKRKRKSFRRTXTKD	900
O	y		899 TEODGEVSAL--GGRAGAGGSPSRKPGGPMGEESPSSCPSPESSEDGPGRSSSPRLVY	956
D	b		901 TEODGEVSALGQGPBARVGPSCRQGPCRPMPGESPESSCGPSPESSEDGPGRSSSPRLVY	960
O	y		957 PFSSPRPPGEPGPPGPELMEDCEKSDTCNPLISGAFSGVINPFSTMGDSRGROYDELPCPC	1016
D	b		961 PFSSPRPPGDPGPPGPLETDEKK--SDTCNPLISGAFSGVINPFSEWGDSRGROYDELPCPC	1019
O	y		1017 APTPSLNTIPLSPSGRRPGRVESRDLAQOLNMLE 1053	
D	b		1020 AAPSLNTIPLSPSGRRSGRDVESRDLAQOLNMLE 1056	
RESULT 2				
ID	08JH78	PRELIMINARY;	PRT; 1186 AA.	
AC	08JH78:			
DT	01-OCT-2002 (TREMBLrel. 22, Created)			
DT	01-OCT-2002 (TREMBLrel. 22, Last sequence update)			
DT	01-MAR-2003 (TREMBLrel. 23, Last annotation update)			
DE	Eig K+ channel.			
GN	ERG.			
OS	Brachydanio rerio (zebrafish) (Danio rerio).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;			
OC	Cyprinidae; Danio.			

Accession	Sequence	Score	Length	Indels	Gaps
NCBI_TaxID=7955;					
RP	SEQUENCE FROM N.A.				
RA	Langheirich U., Wagner T.;				
RT	"Zerg Function in zebrafish."				
RL	Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.				
DR	EMBL; AF532865; AAM95975.1; -				
DR	InterPro: IPR000595; CNMP_binding.				
DR	InterPro: IPR003967; Erg_channel.				
DR	InterPro: IPR005821; Ion_trans.				
DR	InterPro: IPR001622; K+channel_pore.				
DR	InterPro: IPR005820; M+channel_nlg.				
DR	Pfam: PF00027; CNMP_binding; 1.				
DR	Pfam: PF00520; Ion_trans; 1.				
DR	Pfam: PF00785; PAC; 1.				
DR	PRINTS: PR01470; ERCHANNEL.				
DR	SMART: SM0100; CNMP; 1.				
DR	PROSITE: PS50042; CNMP_BINDING_3; 1.				
DR	Ionic channel; Transmembrane.				
DR	SEQUENCE 1186 AA; 132279 MW; BCF3F2351ABBD3 CRC64;				
QY	Query Match	53.3%; Score 3741; DB 13; Length 1186;			
Db	Best Local Similarity 57.5%; Pred. No. 1.2e-235;				
QY	Matches 717; Conservative 121; Mismatches 252; Indels 158; Gaps 36;				
QY	1 MPVARGHVAPONTEFLDTIRKFEQSRKEFIITANAVENCAYIYCNDSGFCSELCGYSRAEYM 60				
Db	1 MPVARGHVALQNTYLTIDTIRKFDQNRKFLANQMKKGIIYCNESGCFQMGFSRAEIM 60				
QY	61 QRPCTCFILGPRPQRRRAADIAQALGAEKKVEIAFYRKDGSCFLLDVVPVKNEDG 120				
Db	61 QQSTCCFLVGPGRMKSALGQALALGSEBKEVIELIYSEGTGRCPLDIDIPVKNEG 120				
QY	121 AVIETFLNFEVVMKKDNY--GSPADHTYHGRPTSMILAPGRAKTFRLKIPALATARES 178				
Db	121 VVIMFILLNFQLLDRSMKKGSLKGRMAN-----SMILAQGRKNIHLMPSL--RVKKQP 172				
QY	179 SVRSGGAGGAGAVVVDVLDLPAPBSSESLADEVTAMDNHVAGLCP---AEERRAL 234				
Db	173 SLDPDHEG-----VVVDY---LQPSHEVALKLDI-----MSPDSCLSKETQAL 214				
QY	235 V--GPGSPRPAQQLPSRAHSLNPDSGSSCLAPTRRSQSCASVSRASADIEAR 292				
Db	215 IQCTPSSCELSPPSRPSRDLREPGC-----LTKSHRESKMSLSLRASSLHDIDGR 267				
QY	293 AGVLPPEPRHASTGAMPLRSLGSLNSTSDSLVRYRTISKIPQITLNFVDLKGDFLASP 352				
Db	268 -----DQMSDLKPSNINSTSDSLMRHRTIGRIQVYISF---GSDRLRP 310				
QY	353 T-SDRETIAP-KIKERTHNVTKE---VYQVLSLADVLPREKLAQPRHRTIILYSPK 407				
Db	311 SPTEIETIAPSKIDRSQNSSEKTYQVYQVLSLGDVLPREKLAQPRHRTIILYSPK 370				
QY	408 AVMWMLILLIYVAVFTYPSAAFLFKETEGEPATEEGCYACORPLAVVDLIVDIMEYDI 467				
Db	371 AVMWMLILLIYVAVFTYPSAAFLFKETEGEPATEEGCYACORPLAVVDLIVDIMEYDI 429				
QY	468 LINFRTTYVYVANEVYSHPRGIAVHYRKGWFLIDMVAIPDDLIFGSGSEB----LIG 522				
Db	430 LINFRTTYVYVNHDEVSNIPARIAQHYRKGWFLIDVIAIPDDLIFRSGSEBQPTTLIG 489				
QY	523 LKFAATLLRLVAVRARKIDRSEYGAAYFLMLCFALAHNLACIYVAILGNNEOHPH--- 579				
Db	490 LKFAATLLRLVAVRARKIDRSEYGAAYFLMLCFALAHNLACIYVAILGNNEOHPH--- 549				
QY	580 -DSRIQGLHNLGDIQGRPYN--SSGLGGBSIXDKYVYVYVYVYVYVYVYVYVYVYVYV 637				
Db	550 GGMKIGWLDNLADIGIKQYNDNSNFSGPSIDKTYVYVYVYVYVYVYVYVYVYVYVYV 609				
QY	638 KIRSIQVWLIGSLMYASIFGVNSAIIQRLYSGTARQYQMLRVRFIFRHOIPNLRL 697				
Db	610 KIRSIQVWLIGSLMYASIFGVNSAIIQRLYSGTARQYQMLRVRFIFRHOIPNLRL 669				

```

Oy 658 EYEOHMASYNGIDMANVLKGPECCLOADICLHNLSLONHCKPEFATGSGCLRALAMK 757
Db 670 EYFOHAPRYNGIDMANVLKGPECCLOADICLHNLSLONHCKPEFATGSGCLRALAMR 729
Oy 758 FKTHAPPGDILVAGDILFALYFISRSIEILKRDYVVALIKNKIDFGEPLNYARPK 817
Db 730 FRTHAPPGDILVHSGDVLTHLYFISRSIEILKRDYVVALIKNKIDFGEPLISLXARPK 789
Oy 818 SNGDVRALTYCDLKHHRDDLLEYLDMPFESDHFWSLIEITFNLKDTNMLPGSPSTEL 877
Db 790 SSADVRALTYCDLKHILKRDLEVLDMYPDFSDNFWMSLIEITFNLKRDYDRIMHPSPSDS 849
Oy 878 EGGSRRQKRLKLSRR-----KTDKDTEDQGEYSALGPRRAGAGPSSRRRPGGWPES 930
Db 850 DCGRRPRHRNRNPLRRNRPDGMDRDMGTVPVOCSPVGNHR-GAIPLSQ-----MDL 902
Oy 931 PSSG-PSSPESEDEGR---GRS---SPLPLVFPSSRRPPGEGPGLMEDECK-SSD 982
Db 903 CSDSPASPLSSEEDMKPLVSGOGMYSLGTEMOERS-----PSAVSLMPSAHSTASA 954
Oy 983 TCNLSGAFS-----GVSNIFFSEWGDSCRQOYELPRCPAPRPSILNIPSSGPRRG 1036
Db 955 MAGPLTGHOQTAPRLNISGYVSYLSIDPKRABEYESQSRSSAYOACYHHHSPCYGDR-N 1013
Oy 1037 DVESRLDALQRLNRLFETRLSADMATVLIQLQROMTIVPAYSAVT-----TRPGPPT 1089
Db 1014 QLARLELQSQLNRLETRMTADINVLIQLQROMAPRPAVYSAVSPRLNHPVRAHPT 1073
Oy 1090 S-----ISPLPV---SPLPLTIDISQVSO-----FMACEELPRGAPR- 1126
Db 1074 SLVYTTAAHNTPTSLQITDASPGKSPDVDSLEKSPDLSISQIHULTVASTJMSMSPETE 1133
Oy 1127 --LPOEG-----TRLSIPGL---GALTSPQ-LHRHGSDP 1157
Db 1134 LSVPSAGFLQPPGLCSLRFPSPLPSDSEGPGLTEGSPETQIRHNSDP 1181

```

QY	121	AYIMEILFEYVMEKDWGSPAHDTNHRGPTSWLAPGRATYFRLKLPALLALTARESSV	180
Db	61	QKPCDCDPLHGPETFRKHADIAQALLGSEERKVEVTVYTHHNGSFICNTHTIIPVKNQEG	120
QY	61	ORPCDCDPLHGPRTFORRAAAGIAQALLGAERKVAIFAYRRKDGSCFLCIVDVVPVKNEDG	120
Db	1	MPVRGRGVAAPONTFLDTIIRKFEEGSKRFTITANARVENCAYIYNDGFCGLGYSRAEYM	60
QY	1	MPVRGRGVAAPONTFLDTIIRKFEEGSKRFTITANARVENCAYIYNDGFCGLGYSRAEYM	60
AC	081V15	PRELIMINARY;	PRF; 732 AA.
DT	01-MAR-2003	(TREMBLrel. 23, Created)	
DT	01-MAR-2003	(TREMBLrel. 23, Last sequence update)	
DT	01-MAR-2003	(TREMBLrel. 23, Last annotation update)	
DE	Similar to potassium voltage-gated channel, subfamily H (eag-related), member 7.		
OS	Homo sapiens (Human).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.		
OX	NCBI_TaxId=9606;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	TISSUE=Brain;		
RA	Strausberg R.;		
DR	Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.		
DR	EMBL; BC035815; AAH35815.1; -		
SO	SEQUENCE 732 AA; 82904 MW; 42E86D3AA12E495D CRC64;		
	Query Match	40.0%; Score 2430; DB 4; Length 732;	
	Best Local Similarity	67.1%; Pred. No. 1,le-174;	
	Matches 490; Conservative 74; Mismatches 134; Indels 32; Gaps 11;		

Db	121	VAMFPIINEFYVTJONENATP----	ERNVPILPIKTVARKEFGFEFPGRLVLTFRKÖSL	175
Oy	181	RSKGAGAGABGAVVVDVLPAPBSSES	LALDEVTAAMDNHVAGLPAEERALVPG--	238
Db	176	POED-----PDVVYID-----	SSKHSDSDVAMKHEFKSPKTESCSPSEADFTKALPSCK	225
Oy	239	SPPSABPOL----	PSPRASHLNPBPASGSGCSCLATRRES	CASVYRASSADDIEMRAG 294
Db	226	SPLVNISGPLHSHSPKQWDRILYPMLO	SSSOLSHSRRESLCSIKRASSVHDIIEF--	G 283
Oy	295	VLPP---PPRASTAMPRLRGLNSTSD	SLVYRRTTSKIPOLTLNVLDGDPELAS	351
Db	284	VHPKPIEFDRHASEEPFNHIKSS	LIGSTSDSNLTKSTINKIPQTLNLSEVYTEKKNS	343
Oy	352	-PTSDRELIAPKIKERTHNVTEKVTQ	VLISGADVLPYEXYKQAPRLHRWTLI	HPKAVW 410
Db	344	PPSSDKTLIAPVKRDNHNVTEKVO	VLISGADVLPYEXYKQAPRLHRWTLI	HPKAVW 403
Oy	411	DWLILLLVYIYRAVFPPIYSAFL	ELKETESGPATCEGYACOP	PLAVDLIYDIMEIVDLIN 470
Db	404	DWLILLLVYIYTAIFPYSAFL	PLNDRRE--OKRREGCYSCSP	PLNVDLIYDIMEIVDLIN 462
Oy	471	FRTYPVVNAANEVYSHGRVAVY	FXGFWELIDWVAALPDLILIFGSGSEE--	LIGILKTA 527
Db	463	FRTYVNVNNEEVYSPAKIAI	HYFGWELIDWVAALPDLILIFGSGSDEFT	LIGILKTA 522
Oy	528	RLRLVYARAKLDARYSEYGA	VLFLMCFALIAHMLACI	YWAIGMNEOPHMDSRIGMIL 587
Db	523	RLRLVYARAKLDARSEYGA	VLMLLMCMCFALIAHMLACI	YWAIGMNEOYVERPYLTDKIGWL 582
Oy	588	NLGDIOGKPYN--SSGLG	BPISKDKYVYALYTFPSS	LTVSGFGVNSPNTNSEKIFTSICVML 646
Db	583	SLGOOIGKRYNDSSSG	PSIKDKYVYALYTFPSS	LTVSGFGVNSPNTNSEKIFTSICVML 642
Oy	647	IGSLMAYSIIFGVVSAIIQ	RLYSGSTARVHTQMLRV	RPIFRHOJPNLRLQRLREYEQHAMS 706
Db	643	IGSLMAYSIIFGVVSAIIQ	RLYSGSTARVHTQMLRV	RPIFRHOJPNLRLQRLREYEQHAWT 702

RESULT 4		
ID	Q18325	
	Q18325	PRELIMINARY; PRT: 791 AA.
AC	Q18325; Q9NG78;	
DT	01-NOV-1996 (TREMBLrel. 01, Created)	
DT	01-OCT-2001 (TREMBLrel. 18, Last sequence update)	
DT	01-MAR-2003 (TREMBLrel. 23, Last annotation update)	
DE	C3OD11.1 protein (HERG-like potassium channel).	
GN	C3OD11.1.	
OS	Caenorhabditis elegans.	
OC	Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidae;	
CC	Rhabditidae; Peloderinae; Caenorhabditis.	
OX	NCBI_TaxID=6239;	
RN	[1]	
RP	SEQUENCE FROM N.A.	
RA	Mortimore B.J.;	
RL	Submitted (Aug-1994) to the EMBL/GenBank/DBJ databases.	
RN	[2]	
RP	SEQUENCE FROM N.A.	
RX	MEDLINE=99069613; PubMed=9851916;	
RA	none;	
RT	"Genome sequence of the nematode C.elegans: A platform for	
RL	investigating biology.";	
RL	Science 282:2012-2018(1998).	
RN	[3]	
RP	SEQUENCE FROM N.A.	
RA	Thomas J.H., Reiner D.J., Newton E.M.;	
RT	"C. elegans unc-103 encodes a HERG-like potassium channel.";	
RL	Submitted (Apr-2000) to the EMBL/GenBank/DBJ databases.	
RN	[4]	

```

RP SEQUENCE FROM N.A.
RA Lloyd C.R. :
RL Submitted (NOV-1994) to the EMBL/GenBank/DBJ databases.
DR EMBL; Z35596; CAAB464.2; -.
DR EMBL; Z46787; CAAB464.2; JOINED.
DR EMBL; AF257518; AAF6899.1; -.
DR EMBL; Z46787; CAC42257.1; -.
DR EMBL; Z35596; CAC42257.1; JOINED.
DR HSSP; P03020; ZCGP.
DR Wormpep; C30D11.1; CE27805.
DR InterPro; IPR000595; cNMP_binding.
DR InterPro; IPR003967; Ecg_channel.
DR InterPro; IPR005821; Ion_trans.
DR InterPro; IPR001622; K+channel_pore.
DR InterPro; IPR005820; M-channel_inj.
DR Pfam; PF00027; cNMP_binding; 1.
DR Pfam; PF00520; Ion_trans; 1.
DR PRINTS; PRO1470; ERGCHANNEL.
DR SMART; SM00100; cNMP; 1.
DR PROSITE; PS50042; cNMP_BINDING_3; 1.
KV Ionic channel; Transmembrane.
SQ SEQUENCE 791 AA; 88518 MW; 5BF65AF56C380BEA CRC64;

```

Query Match	31.9%	Score 1941.5	DB 5	Length 791
Best Local Similarity	52.9%	Pred. No. 1e+137		
Matches 417; Conservative	93	Mismatches 177	Indels 101	Gaps 17

QY	376	OVLISGADVLEBYKLOARIRHWTLLHSPFAVMDYLLLVITVAFTPSAFLKE	435
Db	49	KVLSIGADVLEBYKLOPRHICTVHSPKAVMDYLLLVITVAFTPSAFLKE	108
QY	436	TEEGPRATEGACOPLAVDLVDIMFVLDILINFRTYYANAE--VSHPRIAVH	492
Db	109	LQD---TAKSRTEPELEIVDLVDIMFVLDIINFRTYYVNEDEACVSDPKIATH	165
QY	493	YFKCMFLDMVAALPFDLLIGSGSE---LIGLKTARLLRYAVARKLDYSEYGAV	549
Db	166	YFKCMFLIDMAAEPFDLLVSTNSDEFTTLLGLKTARLLRYAVARKLDYSEYGAV	225
QY	550	LELLMCPTALLAHMLACIMYAIAGNNEOPHMSRIGMLNHLGQIKCPYKSSG---LGAP	605
Db	226	LLLLMAFPALLAHMLACIMYAIAGSELSHKE--YWLHLQDSQLOAPYTSTGTLPTGSP	283
QY	606	SIKQKYVATLYFTSSLLTSVGFGNVSPNTNSEKIFISICMVLGSLMTASIFGNVAIIOR	665
Db	284	TLKSRVYSLXFTLLSTIRSIGFGNVSATTDSEKIPTIIMILIGSLMTASVFGNVAIIOR	343
QY	666	LYSGTARVHQMLRREFIRFHQIPNPLRQRLKEEFGOFANSTNGIDMAVLYKGPPECQ	725
Db	344	LYSGTARVHTMSRIRREFIRFHQIPNPLRQRLKEEFGOFANSTNGIDMVLTKGPDECQ	403
QY	726	ADICLHLNRLSLQHCPRFGATKGCIRALAMKFTTHAPRGDTLHAGDLATLAFISRG	785
Db	404	ADICLHLNRLNLSOGCAAFAGSTPGCLRLSNRFRTHSPRGDTLVHRGDLGLGYFIANG	463
QY	786	SEIILRGD-VVVALIGKNDIFGEPLINTLYARPCKSGNDVRYALTYCDLHKITHRDILLEVDM	844
Db	464	SVELLNDNFTWGLIKGKDIFGENPLLYDEVCKSSCNRYALTYCDLHKITHRDILLEVDM	523
QY	845	YPERSDHWSLEITFNLRDTNMIPGSGSTLEBEGFSRQKRKLSFRRTKDDLEQPG-	903
Db	524	YPERAETFCCKMLTYTNLRD-----DAQSLRKRFNRHKLRLRSSXNMKRYTTPPDG	575
QY	904	-----EVALGPRAGAGSSRRCGPGMGESPSSGSPSESEDECGRSSPRLRY	956
Db	576	DHGNAAVRRSAESVSRCDSNPIDRQSG-----SSSSRCSFPHALATPNTSEARPL	628
QY	957	PFSRPRGEPGPGHPLMEDCEKSSDTCNPLSGAFSGVSNIFSFGDSKGRQYQELP---	1011
Db	629	---LRSTNNHHEBDALFDDLR-----AN-----ARGNTVMTSPYVA	662
QY	1014	-RCPAPPEGLNIPUSSGRRPRGD-----VESRLDALQROLNRLLETLASDMATV	1066

D6 663 GNSVSPPTAIHNDGISHSQSLSDRSDXYERRANFNGRLESIESQEMRMQNKRFSNMDMETL 722

OY 1064 LOLLQROMTL-----VPAY--SAVTTPGPPTSTISPLPVS-----PL 1100
:::|::|::|::|::|

Db 723 IKLVKESDIIINNNGSSNEEPNARYRPNNYYISSAIRLEPGGGGVDDMRVSRLSHSHEPT 782

OY 1101 PTVLRDLNL 1108
||::|

Db 783 PTQETDTI 790

RESULT 5
ID 002497 PRELIMITARY; PRF; 855 AA.
AC 002497;
DT 01-JUL-1997 (TREMBLrel. 04, Created)
DT 01-JUL-1997 (TREMBLrel. 04, Last sequence update)
DT 01-OCT-2002 (TREMBLrel. 22, Last annotation update)
DE SET protein (GH12335P).
GN SET OR ERG OR CG3182.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephyrdoidea; Drosophilidae; Drosophila.
OX NCBI_TaxId=7227;

[1]
RN SEQUENCE FROM N.A.
RP STRAIN-BERKELEY;
RC MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celiker S.E., Holt R.W., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.A., Hoskins R.A., Galie R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Aghayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Bailes R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Bernan B.P., Bhattacharya D., Bolshakov S.,
RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Brotler P.,
RA Butts K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cavley S., Dahlke C., Daventport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Douc P.L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferriz C., Ferlita S., Fleischmann W.,
RA Foster C., Gabriellian A.E., Garq N.S., Gelbart W.M., Glasser K.,
RA Gloddek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Iibegawa C.,
RA Jallali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Keithum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA LaSko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Mishina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusken D.R., Pacלב J.M., Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheefer F., Shen H.,
RA Shue B.C., Sliden-Klamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svrtkas R., Tector C., Turner R., Venter E., Wang A., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zavari J.S., Zhao M., Zhu D., Zhao Q., Zheng L.,
RA Zhong X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
GBbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster."
RL Sequence 287:2185-2195(2000).

RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-CANTON-S;
RX MEDLINE=97178301; Pubmed=8994042;
TT Titus S.A., Warmke J.W., Ganetzky B.;
RT "The Drosophila erg K+ channel polypeptide is encoded by the seizure

RT locus."
RL J. Neurosci. 17:875-881(1997).
[3]
RP SEQUENCE FROM N.A.
RX MEDLINE=97178302; PubMed=8994043;
RA Wang X.J., Reynolds E.R., Deak P., Hall L.M.;
RT "The seizure locus encodes the Drosophila homolog of the HERG
potassium channel."
RL J. Neurosci. 17:882-890(1997).
[4]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkeley;
RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA Champe M., Chavez C., Dorsett V., Farfan D., Frise E., George R.,
RA Gonzalez M., Guarin H., Li P., Liao G., Miranda A., Mungall C.J.,
RA Nunoo J., Paclab J., Paragas V., Park S., Phouanavong S., Wan K.,
RA Yu C., Lewis S.E., Rubin G.M., Celniker S.;
RL Submitted (Oct-2001) to the EMBL/Genbank/DBJ databases.
DR EMBL; AE003462; AAF47148.1; -
DR EMBL; U42204; AAB50956.1; -
DR EMBL; U36925; AAB50936.1; -
DR EMBL; AY058350; AAL13579.1; -
DR FlyBase; FBgn0003353; sel.
DR InterPro; IPR000595; CNMP_binding.
DR InterPro; IPR003967; Erg_channel.
DR InterPro; IPR005821; Ion_trans.
DR InterPro; IPR001622; K_channel_pore.
DR InterPro; IPR005820; M_channel_nlg.
DR Pfam; PF00027; CNMP_binding.1.
DR Pfam; PF00520; Ion_trans.1.
DR PRINTS; PRO1470; ERCHANNEL.
DR SMART; SMO0100; CNMP.1.
DR PROSITE; PSS0042; CNMP_BINDING_3; 1.
KW Ionic channel; Transmembrane.
SQ SEQUENCE 855 AA; 97636 MW; A7251AB788CF3599 CRC64;

Query Match 30.7%; Score 1865.5; DB 5; Length 855;
Best Local Similarity 68.6%; Pred. No. 6.4e-132;
Matches 359; Conservative 64; Mismatches 73; Indels 27; Gaps 7;

QY 355 DRE-TAPKIKERHNATEKYOVLSGADVLPYKIQAPRIHWTLLHSPEKAVDWL 413
DB 270 DKEALLGSKSEPKQDNDMIT--SLGNTLLDQKLNYYHKWTLHSPKFAVMDWI 326
QY 414 ILLVITYAVTPSPAFILKETEGEPATCEGYA-----COPLAVVDLIYDIMEIV 465
DB 327 ILLVMTATITPTVAFLGEOD-----IQRRSKYINSDPIYITDLDVDTFIV 377
QY 466 DILINFRTYVNAEEVYSHPGRIAVHYFKGMFLIDMVAIPFDLLIFGSGSEE--LIG 522
DB 378 DIINFRTYFNQSDDEVYSHGRIAVHYLSGMFLIDVAAVFPDLLVGSOTDDETLTLLIG 437
QY 523 LLKTAARLLRYVARAKIDRSEYGAVALFLIMCTPALLAHMLACIWAIGMEOPHMDSR 582
DB 438 LLKTAARLLRYVARAKIDRSEYGAVALFLIMCTPALLAHMLACIWAIGMEOPHMDSR 497
QY 583 IGMHNLGDOIGRKVYSSGSGGSPISKDYATALYFTSSLTSGVGVNVPNTNEKEFESI 642
DB 498 IGMHNLGDOIGRKVYFNKRTGGSPISKDYATALYFTSSLTSGVGVNVPNTNEKEFESI 557
QY 643 CVMILGSLMVASIFGNVSAIIQRLYSTARYHTOMLVNREFIRPHQIPNPLRQRLSEYFQ 702
DB 558 CVMILGSLMVASIFGNVSAIIQRLYSTARYHTOMLVNREFIRPHQIPNPLRQRLSEYFQ 617
QY 703 HAMTYTGIDMNAVLKGPPECLQADICLHNRSLQHCCKPRGATKGCCLRALAMKEFTTH 762
DB 618 HAMTYTGIDMNAVLKGPPECLQADICLHNRSLQHCCKPRGATKGCCLRALAMKEFTTH 677
QY 763 APPGDTLVHAGDILLTALYFISRSSEIELR-GDVVVALILGKNDIGEPLNTLYARPKSGND 821
DB 678 APPGDTLVHAGDILLTALYFISRSSEIELR-GDVVVALILGKNDIGEPLNTLYARPKSGND 821
QY 822 VRALTYCDLHKIRHDDLEVLDMYPEFSDFHWSLSLETFNLRD 864

DB 736 VRALTYCDLHKIRHDDLEVLDMYPEFSDFHWSLSLETFNLRD 778

RESULT 6
Q8C782 PRELIMINARY; PRT; 515 AA.
AC Q8C782;
ID Q8C782;
DT 01-MAR-2003 (Tremblrel. 23, Created)
DT 01-MAR-2003 (Tremblrel. 23, Last sequence update)
DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
DE Inferred: Mus musculus potassium channel protein.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Heart;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
Nature 420:563-573(2002)."
RL EMBL; AK052366; BAC34958.1; -
DR EMBL; AK052366; BAC34958.1; -
SQ SEQUENCE 515 AA; 58246 MW; 113FC2274308E3FE CRC64;

Query Match 25.2%; Score 1530; DB 11; Length 515;
Best Local Similarity 59.4%; Pred. No. 6.6e-107;
Matches 315; Conservative 64; Mismatches 123; Indels 28; Gaps 9;

QY 1 MPVRGHVAPQNTFLDITIRKEGOSRRTIANNRVENCAYVNGDPCELCYGRAREVM 60
DB 1 MPVRGHVAPQNTFLDITIRKEGOSRRTIANNRVENCAYVNGDPCELCYGRAREVM 60
QY 61 QRPCTCFELHPRTQRRAAQIAQALLGAERKEVEAFYRRKSGSCFLVDVYVVKRNDG 120
DB 61 QRPCTCFELHPRTQRRAAQIAQALLGAERKEVEAFYRRKSGSCFLVDVYVVKRNDG 120
QY 121 AVIFILINFEVYMERKDWGSPANDTNRGPTSWLADGRATFKLTPALIALTARSSV 180
DB 121 VAMFIIINFEVYMERKDWGSPANDTNRGPTSWLADGRATFKLTPALIALTARSSV 175
QY 181 RSGGAGAGAGAVVVDLTPARPSSSESLADEVTAMDNVAGLGAERERALVGG-- 238
DB 176 PQED-----PDVVVID-----SKHSDSVAMKHFKSPYKESCPSEADDTKALIDPSQC 225
QY 239 SPPRSAPGOL----PSPRAHSLNPDAAGSSCSLARTSRSCASVRRASSADIDEARAG 294
DB 226 SPLVNIISGPDHSSPKQMDRLIYDMLOSSQLTHSRRESLCSIRRASSVHDIIEG--S 283
QY 295 VLPP--PBRHASTGAMHPLRSGLLNSTSDSLVRYRTISKIPQITLNFVDLKGDPPLAS 351
DB 284 VHPKNIIFRDRHASSGPNPHIKSSILGSTSDSNLKNKYSTINKIPQITLNFSDVTKETKNTS 343
QY 352 -PTSDRRLIAPKIKERHNATEKYOVLSGADVLPYKIQAPRIHWTLLHSPEKAVW 410
DB 344 PPSSDKTILAKVKEKHNATEKYOVLSGADVLPYKIQAPRIHWTLLHSPEKAVW 403
QY 411 DWLILLVITYAVTPSPAFILKETEGEPATCEGYACOPPLAVAVDLDVIMFVDILIN 470
DB 404 DWLILLVITYAVTPSPAFILKETEGEPATCEGYACOPPLAVAVDLDVIMFVDILIN 462
QY 471 FRTTYVNAEEVYSHPGRIAVHYFKGMFLIDMVAIPFDLLIFGSGSEEL 520
DB 463 FRTTYVNAEEVYSHPGRIAVHYFKGMFLIDMVAIPFDLLIFGSGSEEL 512
RESULT 7
Q8BX82 PRELIMINARY; PRT; 1097 AA.
ID Q8BX82;
AC Q8BX82;

```

DT 01-MAR-2003 (Tremblrel. 23, Created)
DT 01-MAR-2003 (Tremblrel. 23, Last sequence update)
DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
DE ELK channel 3 (Fragment).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Head;
RC MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
DR EMBL: AK048629; BAC33401.1; -.
FT NON_TER 1 122704 MW; D7E37630C1E93B5F CRC64;
SQ SEQUENCE 1097 AA; 122704 MW; D7E37630C1E93B5F CRC64;

Query Match 25.2%; Score 1530; DB 11; Length 1097;
Best Local Similarity 32.0%; Pred. No. 2,1e-106;
Matches 365; Conservative 163; Mismatches 322; Indels 290; Gaps 18;

QY 6 GHVAPQNTFLDTIIRKFGSGSRKFIITANARV-ENCAYIYNDGFCGLGYSRAEYVQRPC 64
   1 GILAPQNTFLDTIATRFDSHNFILANQAVAKGFPYIYCSDFCELAGFARTEVVQKSC 60
QY 65 TQDFLHGPTQRRAAQIQAOLGAERVELAFRKRKQSCFLCYDVVPVKNEDGAVIM 124
   61 SCKFLPGVETNEQMLQMLQKLEEKVEFEGELTFKKKNAPFWCLDIYVINKEGDVAL 120
QY 125 FILNEFVWEKMDVSGSPAHDTNHRGPTSWLAPGAKTFRLKLPALLATARESSVRS6G 184
   121 FLASFK----- 126
QY 185 AGGAGAPGAVVVDLTTPAAPSESLADEVTAMDNHVAQLCPAERRALVPGSPPRS 244
   127 -----DIDDTYKVKITSEDKKEDETRGRSR----- 150
QY 245 PGQLSPRAHSLNPASGSSCLARTRSRESCASVRRASSADDIEMRAGVLP PPPRHS 304
   151 -----AGSHEDSARRSR----- 163
QY 305 TGAMHPLRSLNSTSDSLVRYRTISKIPQITLNFVDLKGDPFLASPTSDREIAPKIK 364
   164 -----AVLXHSGLHQREKKLKNNVFVDK----- 192
QY 365 ERTHNVTKEVTOVLSGADVLDPEYKIQADRIHRWTILHSPKAVWDMLILLVITYAVF 424
   193 -----APEYKASDAKSKSFILHSTFKAGDWMLILLATFYVAVT 233
QY 425 TPYSAFLIKETEBGPATECGYACQPLAVVDLIYDIMEIVLILNFRITYNANEVVS 484
   234 VYNNCFIGNEDLSTTRST-----TVSDIAVEILFIDILLNFRITYVSKSGVIF 284
QY 485 HPGRIAVHYFKGFILDMVAIPDL-IFGSGSEELIGLKTARLLRIVRARKDRS 543
   285 EARSICIHVYTWFTIIDLAIPDLIAFNVTYVSLVLLKTVRLIRLRLKQIDRS 344
QY 544 EYGAAVLFLMCTFALLIAHMLACIYVAGNMEOPH--MDSRIQWLHNLGDIQGR-YNS 599
   345 QHSTIVLTLTMSFALLAHMWACIYIIGKMEREDNSLLKWEVGHLELGRLESPPYGN 404
QY 600 SGLGSPSTKDKVYVTLTFEFSSTVGCGNVSPNTNSEKISICVWLIGSLMYASIFGV 659
   405 NTLGSPSTRSAYIALVFTLSSTVCGNVSANDAEIKFISICVWLIGALHVALVFGV 464
QY 660 SAIIRLYSGATRYTOMLRVREFIRFOIIPNPLRQRLREYFOHMASTYNGIDMAVLKG 719
   465 TAIIRMRSRMSLHYTRKDKLDFIRVHHLPOOLQRMLEYFOTIWSVANGIDSNELKD 524

```

```

QY 720 PPECIQAIDICHLNLSLLOHCKPFRGATRGCLRALAMKFTTHAPGDTLVAGDILTL 779
   525 PFDELRSITDTHMLDKREIIO-LSLFECSARGCRSLSLHKTYSFCAPEGVLLKQDALOI 583
QY 780 FYTSGSIEILRGDVVALITGKNDFJGPELNLVYARSGKSGVRLATYCDLKKIRDDL 839
   584 YVYCSGSMKVLKDSMVALITGKDLIGANLSIKDOYIKTNADYKALITYCDLCKILKIF 643
QY 840 EYLDWYPERSDHFWSSL--EITFNLRD-----TNMIPGSPSTLEGGFSRQRKR 887
   644 EYLGIXPEYAHKFVEDIQHDLTYNLREGHSDVYIRLSKNSKSVQSAEPKNGSIN--KR 700
QY 888 KLSFRKRDYKTEQCEVSALOPGRAGAPSSRGAPG----- 925
   701 IPSIYDEDEEEVEEETTSLSPIYTGSSVSHSKTGSNKTYLGLSLKQLASGVTPFHS 760
QY 926 PWGESPPSGSPSESEDEGPRSSPLLVFPSSPRPGEPGPELMEDEKSDICN 985
   761 PIRVSANSKPKQKQADPNHGRKKKLNKVLSSLGSACTPBLSPRYDIE----- 813
QY 986 PLSGAFSGVSNIFSEWGSRGROYELPCPAPFPSLNIPLSSGRRRPGVESRLDA- 1044
   814 -----DGNMNETQTFDQSGEQIRPEPRISPL-----ASEBIGLA 849
QY 1045 -----LQRLNLE-----TRLSADMATVLDLQRLQMTLVPPAYSAYTTQGP 1086
   850 FLFIRAETKQIOLINKINSEVTTLTQEVSLGDRHMSIMOLLNLTLSPOQPSOFCSLHPR 909
QY RESULT 8
   ID Q8CC38 PRELIMINARY; PRT; 522 AA.
AC Q8CC38;
DT 01-MAR-2003 (Tremblrel. 23, Created)
DT 01-MAR-2003 (Tremblrel. 23, Last sequence update)
DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
DE Inferred: Mus musculus potassium channel protein.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Diencephalon;
RC MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
DR EMBL: AK034003; BAC28543.1; -.
SQ SEQUENCE 522 AA; 59030 MW; D687719147B84D38 CRC64;

Query Match 24.9%; Score 1516.5; DB 11; Length 522;
Best Local Similarity 58.7%; Pred. No. 7,1e-106;
Matches 315; Conservative 64; Mismatches 123; Indels 35; Gaps 10;

QY 1 MPVRGHVAPQNTFLDTIIRKFGSGSRKFIITANARVENCAYIYNDGFCGLGYSRAEY 60
   1 MPVRGHVAPQNTFLDTIIRKFGSGSRKFIITANARVENCAYIYNDGFCGLGYSRAEY 60
QY 61 QRPCTCDLHGPTQRRAAQIQAOLGAERKEVIAFYRKDSCFLCLVDVVPVKNEDG 120
   61 QRPCTCDLHGPTQRRAAQIQAOLGAERKEVIAFYRKDSCFLCLVDVVPVKNEDG 120
QY 121 AVIMFILNEFVWEKMDVSGSPAHDTNHRGPTSWLAPGAKTFRLKLPALLATARESS 180
   121 VAMFIIINEFYVTDENATP-----ERYNPILPVKYNRKLFGKFPGLRLVLYRKQSL 175
QY 181 RSGAGAGAPGAVVVDLTTPAAPSESLADEVTAMDNHVAQLCPAERRALVPGSP-- 238
   176 PQED-----PDVYVID-----SSKHSDSVAMKHKFKPKESCSFSEADDTKALIQPSOC 225

```


Query Match 23.9%; Score 1453.5; DB 5; Length 1174;
 Best Local Similarity 30.1%; Pred. No. 1.4e-100;
 Matches 375; Conservative 173; Mismatches 314; Indels 383; Gaps 32;

```

QY 4 RGHVAPONTFLDTIIRKFEQO-SRKFTIARVENCAYIYCNDFGCELCGSRRAEVMOR 62
DB 5 RRGVLAPONTFLDTIIRKFEQO-SRKFTIARVENCAYIYCNDFGCELCGSRRAEVMOR 64
QY 63 PC-TCDLHGPRTORRAAQAIALGAERKVEIAFYRKDG--SCFL----- 107
DB 65 SCRVCGFMYGELTKETVGRLEYTLLENQOQOPELLYKKNLCCGCLSPFGKAQTOE 124
QY 108 ----CLVDVYVKNEDGAVIMFILNEEYVMEKDMGSPADHTNHGPTSWLAPGRAKTE 163
DB 125 TPLMLLOVAPIRNRERDLVIFLL-----TF 150
QY 164 RLKLPALLALTARESSVRSRGAGAGAPGVVVVDLPPAPSSSIALDEVTAMDHNVA 223
DB 151 R-----DITLAKOP1- 160
QY 224 GLGPAEERRALVPGSPPRSAPGOLPPRAHSLNDASGSSCLAR--TRSHESCAVVR 281
DB 161 ---DSEDTRKGYLG-----LSKFAKLARSVYRSNGQFSAHL-- 191
QY 282 ASSADDIEMRAGVLPPEPRRSTGAMHPLNSGLNSTSDSLVRYRTISKIPQITLNFV 341
DB 192 ----- 191
QY 342 DLKDPFLASPTSDREITAPKIKERTHNTEKVTOLSGADVLEPYKLOAPRIRHWTL 401
DB 192 ----PILKDPTRKO-----SNLHMMSISADIMQYKREAPKTPPHIL 230
QY 402 HSPFKAVMDLILLVITYAVETPYSAAFILKETEEGPRATECGACQPLAVVDLIVDI 461
DB 231 HCAKAKAIWDWYIILCTEYTAIMVYNYAFKNKTSED-----VSLIVDSIYDV 279
QY 462 MEIVVILINERTYVNAMEVYSHRGRTAVHFKMPLIDMYAALPFDLLIGS----- 515
DB 280 IFDFIVINFTEFVPGGEVYSDPKVIRMNLIKWFIIIDLSCLPYD--VFNADRDDED 337
QY 516 GSEELIGLTKARLRLVVARVARDRYSEYGAVALFLMCPFALLAHMLACTIWAYIGNME 575
DB 338 GIGSLFSAALKVYRLRLGRVYARKLDRIEYGAAMLILLCFMVAHMLACIWTISIGSD 397
QY 576 QPHMDSRIGLHNLGDIQKPY-----NSSG--LGPISIKKYVYALYTFESSLTSYGF 627
DB 398 ADN-GIOYSWLMKLANVQSPSYTMSNDTGPELVNCGPSRKMVYALYFTWTCMTSYGF 456
QY 628 GNVSPNTNSEKIFSCVMLIGSLMAYSTEGNVSATIQRLYSTAYHYOMLRYREFIRH 687
DB 457 GNVAAETNEKEVFTICMMLIALATLFGHTVTTIIQMTSATYAKYHMLNNVREFMKLH 516
QY 688 QIPNRLORLEVEYFOHANSYTGIDMNAVLKGFPECLADICLJHNRSLJQHCXFRGAT 747
DB 517 EYPRKALSRVMDYVYVSTWAMTKGDLTEKVLNCPDMKADICVHLNKRKYFNHPRFRLAS 576
QY 748 KGCIALAMKRTTHAPPGDILVAAGDILTALEYISKCSIELTKRDVVAALIGKNDIFGE 807
DB 577 DGCLFALAMHFMMSHAPGDLIYHGESIDSLCFIVTSLEVIQDEYVAALIGKGDVFGD 636
QY 808 PLNLVAPRGKSGNDVRALTYGDLHKIHRDDELLEVLDMPERPDHWSLEITFNLRDNNM 867
DB 637 QPMKDSAVGOSANVRALTYCDLHAIKRDKLEVLIDFYSAPANSFARNLVLYTNHRLH 696
QY 868 I-----PGSP-----GSTLEGG----- 880
DB 697 FRKXADVAREKEIARRKNEPOLPONOHLYRKIFSKRRTRTPVOVAGSKELVGSQSDV 756
QY 881 -----FSHQRKRLSFRRTDKDT-EDGPEYSAIOPGRAGAPSSRGPRGPW 927
DB 757 EKGDEVEYRTKYVFPAPKRLQASQATLARDQFIDEGEVDSSPPSR-----DSRVYIEGAA 811
QY 928 GESFSSSGSPSSSESD-----GGRSSSPILRLVFPSSPRPGEPGGEPLMEDCE 978

```

Query Match 23.4%; Score 1423; DB 5; Length 956;
 Best Local Similarity 30.1%; Pred. No. 2.1e-98;
 Matches 357; Conservative 166; Mismatches 323; Indels 342; Gaps 26;

```

DB 812 VSSATVGPSPVATTSAAAGAVSGGPGSGGTVAIVYTKADRNLTALEREIOIEMASSRA 871
QY 979 KSSDFCNPLSGAFSGVSNIFSPWGSRGQYQELRCAPAFSLINITPLSSGRRPRGPV 1038
DB 872 TTSDTYD-----TGL-----RET-----PTLQORDLIATVLDMKVDV 904
QY 1039 ESRDLAOLNRLNLETRLSADMAVYLOLQORMTLVPPA-----YSAVTPGP--- 1086
DB 905 RLELQRMQORIGRIDLLG-----ELVKR--LAPGAGSGGNAPDNSSGQTTTGDELC 954
QY 1087 -----GPTSTPLPVSPPLTLTDSISOVSQPMACEBELPGCA 1124
DB 955 AGCGAGGGTPTTQAPPTSAY-TSPVDVYITISS-----PGA 990

RESULT 12
Q9XYX7
AC Q9XYX7 PRELIMINARY; PRT; 956 AA.
DT 01-NOV-1999 (Tremblrel. 12, Created)
DT 01-NOV-1999 (Tremblrel. 12, Last sequence update)
DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
DE EAG K+ channel EGL-2.
GN EGL-2.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxId=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=N2;
RA Welshancker D., Wei A., Salkoff L., Thomas J.H.;
RT "Regulation of cell excitability by egl-2, an EAG-like K+ channel, blocked by a triacyclic antidepressant in Caenorhabditis elegans.";
RL Submitted (Feb-1999) to the EMBL/Genbank/DBJ databases.
DR EMBL: AF130443; AAD28468.1.
DR InterPro: IPR005955; cNMP_binding.
DR InterPro: IPR005821; Ion_trans.
DR InterPro: IPR001622; K+channel_pore.
DR InterPro: IPR005820; M+channel_nlg.
DR InterPro: IPR001610; PAC.
DR InterPro: IPR007000; PAS-assoc_C.
DR InterPro: IPR000014; PAS_domain.
DR InterPro: IPR002016; Peroxidase.
DR Pfam: PF00027; cNMP_binding.
DR Pfam: PF00520; Ion_trans.
DR Pfam: PF00785; PAC.
DR SMART: SM00100; cNMP.
DR SMART: SM00086; PAC.
DR SMART: SM00091; PAS.
DR TIGRFAMs: TIGR00229; sensory_box.
DR PROSITE: PS50042; cNMP_BINDING_3.
DR PROSITE: PS50113; PAC.
DR PROSITE: PS50112; PAS.
DR PROSITE: PS00435; PEROXIDASE_1.
DR PROSITE: PS00435; PEROXIDASE_1.
KW Ionic channel; Transmembrane.
SQ SEQUENCE 956 AA; 108091 MW; E998B7F212805273 CRC64;

```

```

Db 121 DAVVLYLCOFK----- 131
Qy 180 VRSGAGAGAGAVVVDVLTLPAPSSSESLALDEVMTAMDNHVGIGPAERRALVPGS 239
Db 132 -----DITP----- 135
Qy 240 PPRAPGOLPSFRAHSLNPDASGSSCSLARTRSRESCASVRRASADDIEMRAGVLP 299
Db 136 -----LKP----- 139
Qy 300 PRHASTGAMHPLRSLNSTSDSLVRYRTISKIPQITLNVYDLKGPPLASPTDREIT 359
Db 140 -----LDENKNGKSLRIQIARIRKSKQOFOIETKDLHKSP----- 176
Qy 360 APKIKERTHNVEKTVOLISGADVLPKYLOAPRIHRTILHSPFRVAMWMLLLVI 419
Db 177 -----GNTSSNFNOVMNLGDMLPQROPTPTSPHILHYSFPTIMWMLTALTE 228
Qy 420 YTAVEPTPSAA-----FLKETECPATECGACOPPLAVVDLIVIMFIVDILINR 472
Db 229 YTAVEPTPSAA-----GGIDSVALMDSIVDIYFADILLNFH 283
Qy 473 TTYVANAEVSSHPRIAVHYFKWFLIDMAIPDLLIEGSGSEELIG---LKTAR 528
Db 284 TTFVGPGEVYIEPVIQNYFKSWFLIDLSCLPYDFYMKRDERIGLSFALKVYR 343
Qy 529 LLRLVRAVRKIDRYSEYGAVALFLMCTFALAHMLACVYAGNME-QPHNDSKI---G 584
Db 344 LLRLGVRAVRKIDRYSEYGAVALFLMCTFALAHMLACVYAGNME-QPHNDSKI---G 584
Qy 585 WLNHNDGIDGKPYN-----SSGLGSPSISKDYVYALYFTFESSLSVSGNVPMTNSEK 638
Db 404 WLMKLSNDLRQVYNIPLSKKTLVGGPSTSAVTSISLVYTGSCMTVGCGNASTMDNK 463
Qy 639 IFSIVMLIGSLMVAISFGNVAISATIQRLYSGTARYHTOMLVRREFIRHQIPNPLRLE 698
Db 464 IFGVCMITISALLAIFGHMTTIIQOMTSSVTRYHEMISNVEFIKLOEIKELAEVY 523
Qy 699 EFGHAWSTYNGIDMNAVLKGFPECLQDICHILHRSLLQHKCPRGATKGLRLAMKF 758
Db 524 DTVSTWMTAKGIDPAKVLGYCPKDMKADICVHLKRNKFNHESCRSLSDGSLRLAMFL 583
Qy 759 KTTNAPGDTLVHAGDILLATYFISRSIEILRGDVVAALGKNDJFGEPL-NLYARPGK 817
Db 584 ELNHAPODLYHTGESVDALMFVYSGSLVYQDDEVAAILGKGVFGEFEKANGSTGQ 643
Qy 818 SNGDVRLTYCDLHRIHDDLEVLDMYPEFSDFWSSLEITFNLRTNMIIPSGSGTEL 877
Db 644 SAANVRALTYSDLHMKIKKDKLMDVLDYKAFANSFARNMTLYNLTH-RMKFRYADVYR 702
Qy 878 ECGFSRQK-----RKLSFRF-----IDKTEQGEVYALGP 910
Db 703 EELDAKKRNEKLLPNHPIRKLLEFRMERHGPRIFPSMPADIEKGLKKTSTLSRI-- 760
Qy 911 GRAGAPSSRGRPGGPMGESSPSSSEDEGPRGSSPLRIYFESSPPGPEPGC 970
Db 761 -----SSLHSMIDETGGGSSVYK-SPKSKPRP----- 788
Qy 971 EPLMEDCEKSDPTCNPLSGAFSGVSNIFSEWG-DSRGROYOLPRCPAPPTSLNIPSS 1029
Db 789 -PLMQRQVDEDLARTS-----WGMDDKREWSLSLNKTEMKSKFPI----- 831
Qy 1030 PGRRRP--GDVESRLDALOR-----QNLNLEPLASDMATVYOL 1066
Db 832 IGERLTIIEQINSRLALLERVLIGNNGANTPSTMPVGSFALSNGNLTLDAAPVARS 891
Qy 1067 L-----OROMTLVPPAYS-----AVTTPGPGPTSTPLP 1096
Db 892 VMSBQHQHPHQRTST-VPLRLRELAGEWEPRIREFTPNRTSSSRVP 938

```

RESULT 13
044164

```

ID 044164 PRELIMINARY; PRT; 956 AA.
AC 044164;
DT 01-JUN-1998 (TREMBLrel. 06, Created)
DT 01-OCT-2001 (TREMBLrel. 18, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE Hypothetical 108.1 kDa protein.
GN F16B3.1.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxId=6239;
RN 11
RP SEQUENCE FROM N.A.
RC STRAIN-Bristol N2;
RC MEDLINE=99069613; PubMed=9851916;
RA None;
RT "Genome sequence of the nematode C. elegans: a platform for
RT investigating biology. The C. elegans Sequencing Consortium.";
RL Science 282:2012-2018(1998).
RN 12
RP SEQUENCE FROM N.A.
RC STRAIN-Bristol N2;
RA Kemp K., Wilson R.;
RT "The sequence of C. elegans cosmid F16B3.";
RL Submitted (DEC-1997) to the EMBL/Genbank/DBJ databases.
RN 13
RP SEQUENCE FROM N.A.
RC STRAIN-Bristol N2;
RA Waterston R.;
RT "Direct Submission.";
RL Submitted (JUN-2001) to the EMBL/Genbank/DBJ databases.
DR EMBL AF036695; AAB8348.2; --
DR WormPep: F16B3.1; CE24905.
DR InterPro: IPR000595; CNMP_binding.
DR InterPro: IPR005821; Ion_trans.
DR InterPro: IPR001622; K+channel_pore.
DR InterPro: IPR005820; M+channel_nlg.
DR InterPro: IPR001610; PAC.
DR InterPro: IPR000700; PAS-assoc_C.
DR InterPro: IPR000014; PAS_domain.
DR InterPro: IPR002016; Peroxidase.
DR Pfam: PF00027; CNMP_binding_1.
DR Pfam: PF00520; Ion_trans_1.
DR Pfam: PF00785; PAC_1.
DR SMART; SM00100; CNMP_1.
DR SMART; SM00086; PAC_1.
DR SMART; SM00091; PAS_1.
DR TIGRFAMs; TIGR00229; sensory_box_1.
DR PROSITE; PS00042; CNMP_BINDING_3; 1.
DR PROSITE; PS00113; PAC_1.
DR PROSITE; PS00112; PAS_1.
DR PROSITE; PS00435; PEROXIDASE_1; 1.
DR Hypothetical protein; Ion channel; Transmembrane.
KW SEQUENCE 956 AA; 108090 MW; E998B7F2128054A8 CRC64;
SQ

```

Query Match 23.4%; Score 1423; DB 5; Length 956;
Best Local Similarity 30.1%; Pred. No. 2,1e-98;
Matches 357; Conservative 166; Mismatches 323; Indels 342; Gaps 26;

```

Qy 1 MPV-RGHVAPONTFLDTIRKFEQOSRKFIITANARVENCAYIYNDGFCELGYSRAEV 59
Db 1 MPVGRGLVAPONTFLENVIRRCNNADTSFILANQVVDYPIYVNCNDSGSKLVGTYRAEI 60
Qy 60 MORPCTDPLHPRKPRQORAAOIAOALLGAERKVEIAYFRKDGSCFLCVYVYVYKND 119
Db 61 MOKPCLAFMGEHEHEVGSLOKMOBALENARTEQAEIGCKKNKPIVLLVHLAIPKHK 120
Qy 120 GAVIMEIINFEVMEKDWGSPAHDTNHRGPTSMILAPRAKTFRLKLPALLATPARESS 179
Db 121 DAVVLYLCOFK----- 131
Qy 180 VRSGAGAGAGAVVVDVLTLPAPSSSESLALDEVMTAMDNHVGIGPAERRALVPGS 239

```

Db 132 -----DITP----- 135
Qy 240 PPSAPQOLSPRAHSLNPDASGSSCSLARTSRSCASVARRASSADDIEAMRAGVLP 299
136 -----LKQP----- 139
Qy 300 PRAHSTGAMHPLNSGLNSTSDSLVRYRTISKIPQITLNFVDIKGDFLASPTSDRELI 359
140 -----LDENNKLSRILOIARIASKOOFNOIETDKLHKSP----- 176
Qy 360 APRIKERTHNVTEKVNOVLSIGADVLEPKLOARIRHWTILHSPKAWMDLILLI 419
177 -----GNTSSNFNOVMNLGGDMLPYQKQETPKSPHILHYSKFTIMDWSILALTF 228
Qy 420 YTAVFPTYSAA-----FLKETEESGPATECGYACOPLAVDLIYDIMEIVILINFR 472
229 YTAFMVFNIAFKNSLRPFILISRENP-----CGGIDVALMDSIYDVFADILLNFH 283
Db 473 TTYVNAMEEYVSHGRJAVHYFKGWLIDMVAIIPDLLIFGSGSEELIG---LKTPAR 528
284 TTFVPGGEVIEPVSIVROYFKSMFLIDLSCLPYDIFYWFKRDDRIGSLFALKVVR 343
Qy 529 LLRIYVARKLDRSEGAVALFLMCTFALIAHMLACIYALGNME-QPHMDSRI---G 584
344 LLRIYVARKLDRSEGAVALFLMCTFALIAHMLACIYALGNME-QPHMDSRI---G 584
Qy 585 WLNHIGDOIGKPYN-----SSGIGSPISDKRYTALYFTFSSLSYVGFNVSPNTSEK 638
404 WMLNLSMDLQNHVYIPLSNKTTLYGSPSRISAYISLSLYTMSCTSYGFGIAANTINEK 463
Qy 639 IFSICVMLIGSLMAYISIFGNVSAIIORLYSGTARYHOMLRREFIRPHQIPNPLKORLE 698
464 IFGVCMMIISALYAAIFGHMTTIIQOMTSYVRYHEMISVREFIKLOEIPKLABERV 523
Qy 699 EYFQHANSYNTGIDMNVLAGFPECLQADICHLNRSILOKCPREGATKCALAMKF 758
524 DYVYSTAMTKGIDTAVALGICPKDKADICVHLNRYFNHSCFRLASDGLSLAMFL 583
Qy 759 KTHHAPPGDTLVHAGDILLALYFISRSIELRGDVVAALIGKNDIGEPL-NLYARPGK 817
584 ELNHAAPGDDLHYHGESVDALMFVSGSLVLEIODEVVAALIGKDVGEDEFWAKNGSTQ 643
Qy 818 SNGDVRLTYCDLHKIRDDILEVDMYPERSDHFWSSLETFNLRDTNMIPGSPSTEL 877
644 SAAVNRALTYSDLMIMIKKLDVLYKAFANSFARNMUTLVNLT-RMFRKRVADYK 702
Qy 878 EGGFSRORK-----RKLSFRRR-----TDKDTBOGGEVSLGP 910
703 EKELDARKNKNEKTLPHDHPRLKLFMRKREHGRIPSPMFADIEKGLKSTELSKI-- 760
Qy 911 GRAGAGPSSRGRPGGPGSPSSPESSEDEGPRSSPLRLVFPSSPRPPEPPGG 970
761 -----SSLHSMIDETGGGSSSYK-SPRSKRKR- 788
Qy 971 EPLMEDEKSDTCNPISGAFSGVSNIFSWG-DSRGROYOELPRCAPTPSLNLPLSS 1029
789 -PLMKROTVEDDALSRYS-----WGMKKDKREWSLSINIKTEMSKRPDI--- 831
Qy 1030 PGRMR-GRVESHDLALOR-----QLNRLETRLASADATYLOL 1066
832 IGERLTIIEQINSFLALIERVLIGNNGANTPSTMPVGSFALSSEGNRLDAAPARS 891
Qy 1067 L-----QROMTLVPPAYS-----AVTTPPGPPTSLPLP 1096
892 VSMSEQHQPWKQRTST-VPLRLLEAGWEPRIREPFPNPTSTSSRPV 938
RESULT 14
ID Q8C035 PRELIMINARY; PRT; 988 AA.
AC Q8C035;
DT 01-MAR-2003 (Tremblrel. 23, Created)
DT 01-MAR-2003 (Tremblrel. 23, Last sequence update)

DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
DE EAG K+ channel.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=olfactory brain;
RX MEDLINE=2254683; PubMed=12466851;
RA THE FANTOM Consortium,
RA the Riken Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs."
RL Nature 420:563-573(2002).
DR EMBL: AK032438; BAC27869.1;
SQ SEQUENCE. 988 AA; 11781 MW; 0EE1F577F5C18BB CRC64;
Query Match 23.18; Score 1403.5; DB 11; Length 988;
Best Local Similarity 28.88; Pred. No. 6.5e-97;
Matches 354; Conservative 177; Mismatches 337; Indels 363; Gaps 24;
Qy 4 RGHVAPONTFLDITIRKFEQSRKFTIANARVENCAYICNGFCGLGYSAEYVQRP 63
Db 5 KRGLVAPONTFLEIVIRSRSESS--FLLGNQVQIDWPFVINSDFCKLSGTHRADVQKS 62
Qy 64 CTQDFELHPTORRAAQIAQALIGAEERKEVEIAFYKRKDSCEFLCTDVVPVKNEDGAVI 123
Db 63 STGCFMNGELLDKTIKRYQTFPNYNSCEVLLYTKNTPWFYMQIAPIRHEKVV 122
Qy 124 MFIINEFVNEKDNVSPADHTNHRGPTSMLABGRKTRLLKLPALLALTARESSVRSG 183
Db 123 LFLCTFEDI--LFLKOPIEDDSYTKG--WTKFAR----- 151
Qy 184 GAGAGAGPAGVAVVDVDTTPAPASESLADEVTAPADNHVAVGLGPAEERALVGPSPRS 243
Db 152 -----LTRLATNSRV-LOQLTPM----- 170
Qy 244 APQOLSPRAHSLNPDASGSSCSLARTSRSCASVARRASSADDIEAMRAGVLP 303
171 ----- 170
Qy 304 STGAMHPLRSGLNSTSDSLVRYRTISKIPQITLNFVDIKGDFLASPTSDRELIAPKI 363
171 ----- 170
Qy 364 KERTHNVTEKTVOLSGADVLEPKLOAPRIHRTILHYSPEKAVMDLILLIYTA 423
171 KETVYHKSRLAEVLOGLSDILPYQKQAPKTRPHIILHCAFTTMDWYILLIFETAI 230
Qy 424 FTYPSAFLKETEESGPATECGYACOPLAVDLIYDIMEIVILINFRTTYVNAMEEV 483
231 MYPYNAVSEFKRONNIA-----WVLDSVDVIFLVDIVLNFHTTFVPGGEV 278
Qy 484 SHGRJAVHYFKGFLDMMVAIIPDLL---IFGSSSEELIGLTKARLLRVARVARK 539
279 SDPLRIMNYLKTFFVIDLSCLPYDIIINAFENVDEGISLFSLLKAVRLRLGRVARK 338
Qy 540 DRYSEGAVALFLMCTFALIAHMLACIYALGNME-----QPHMDSRIQLHNDQ 592
339 DHTIEGAAVALVLYCVGLVAHMLACIYALGIDYEIVDEVTNTQIDS--WLYQALS 395
Qy 593 IGKPY---GSPSIXDKVYVTLALYFTFSSLSYVGFNVSPNTNSEKISICVMI 647
396 IGTFYRYNTSAGIWEGPSNDSLVSLLYFTMNSLITIGGNIAPTDVEKMPSVAMMY 455
Qy 648 GSLMAYISIFGNVSAIIORLYSGTARYHOMLRREFIRPHQIPNPLKORLEEFQAHV 707
456 GSILYATIFGNVTTIIFOOMYANTMYRHEMNNVADFLKLYOVPGLSERWADYIVSTWS 515
Qy 708 TNGIDMNAVLKGPPECLQADICHLNRSILOKCPFGATKGCRLALMEKTHHAPPG 767
516 SKGIDTEKVLISICKDKRADICVHLNRYFNHSCFRLASDGLRLALAVEFOTIHCA 575

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 23, 2003, 17:35:40 ; Search time 37 Seconds

(without alignments)
1325.359 Million cell updates/sec

Title: US-10-000-151b-3

Perfect score: 6079
Sequence: 1 MPVRGHVAPQNTFLDTIIR.....GOLGALTSPFLRHSGSDPGS 1159Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

Issued_Patents_AA:*
1: /cgn2_6/ptodata/2/1aa/5A.COMB.pep:*
2: /cgn2_6/ptodata/2/1aa/5B.COMB.pep:*
3: /cgn2_6/ptodata/2/1aa/6A.COMB.pep:*
4: /cgn2_6/ptodata/2/1aa/6B.COMB.pep:*
5: /cgn2_6/ptodata/2/1aa/PTCTUS.COMB.pep:*
6: /cgn2_6/ptodata/2/1aa/backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	6079	100.0	1159	2 US-08-956-242-13	Sequence 13, Appl
2	6079	100.0	1159	3 US-09-351-215-13	Sequence 13, Appl
3	6079	100.0	1159	3 US-09-226-012-2	Sequence 2, Appl
4	6079	100.0	1159	3 US-09-226-012-4	Sequence 4, Appl
5	6079	100.0	1159	4 US-09-358-383C-10	Sequence 10, Appl
6	2436.5	40.1	888	2 US-08-956-242-4	Sequence 4, Appl
7	2436.5	40.1	888	3 US-09-351-215-4	Sequence 4, Appl
8	1653	27.2	626	2 US-08-956-242-2	Sequence 2, Appl
9	1653	27.2	626	3 US-09-351-215-2	Sequence 2, Appl
10	1636	26.9	1017	4 US-09-600-776-6	Sequence 6, Appl
11	1568	25.8	1107	4 US-09-358-383C-16	Sequence 16, Appl
12	1551	25.5	1102	4 US-09-358-383C-36	Sequence 36, Appl
13	1541.5	25.4	1083	4 US-09-343-494-1	Sequence 1, Appl
14	1340.5	25.3	1083	4 US-09-600-776-2	Sequence 2, Appl
15	1340.5	25.3	1083	4 US-09-358-383C-2	Sequence 2, Appl
16	1536	25.3	1082	4 US-09-336-643A-20	Sequence 20, Appl
17	1485.5	24.4	1284	4 US-09-343-494-9	Sequence 9, Appl
18	1485.5	24.4	1284	4 US-09-358-383C-11	Sequence 11, Appl
19	1415	23.3	962	4 US-09-614-480-9	Sequence 9, Appl
20	1399.5	23.0	988	4 US-09-614-480-2	Sequence 2, Appl
21	1190	19.6	542	4 US-09-358-383C-5	Sequence 5, Appl
22	716.5	11.8	241	4 US-09-358-383C-23	Sequence 23, Appl
23	565	9.3	181	4 US-09-358-383C-26	Sequence 26, Appl
24	520	8.6	749	4 US-08-997-685A-10	Sequence 10, Appl
25	515.5	8.5	910	4 US-08-997-685A-2	Sequence 2, Appl
26	470	7.7	528	4 US-08-997-685A-4	Sequence 4, Appl
27	460.5	7.6	290	4 US-09-358-383C-8	Sequence 8, Appl

28	460	7.6	597	4 US-08-997-685A-12	Sequence 12, Appl
29	445	7.3	506	4 US-08-997-685A-6	Sequence 6, Appl
30	436.5	7.2	319	4 US-09-358-383C-22	Sequence 22, Appl
31	351.5	5.8	180	4 US-09-358-383C-25	Sequence 25, Appl
32	334	5.5	189	4 US-09-358-383C-25	Sequence 25, Appl
33	315.5	5.2	126	4 US-08-997-685A-59	Sequence 59, Appl
34	303.5	5.0	58	4 US-09-358-383C-29	Sequence 29, Appl
35	248	4.1	52	4 US-09-358-383C-35	Sequence 35, Appl
36	233.5	3.8	89	4 US-09-358-383C-30	Sequence 30, Appl
37	232.5	3.8	170	4 US-09-358-383C-27	Sequence 27, Appl
38	216	3.6	75	4 US-09-358-383C-28	Sequence 28, Appl
39	214	3.5	90	4 US-09-358-383C-31	Sequence 31, Appl
40	196.5	3.2	2353	4 US-08-984-709A-50	Sequence 50, Appl
41	184.5	3.0	1180	3 US-08-660-148-2	Sequence 2, Appl
42	182	3.0	2273	4 US-09-426-998-5	Sequence 5, Appl
43	181.5	3.0	1180	1 US-08-072-574-8	Sequence 8, Appl
44	179.5	3.0	1180	1 US-08-466-270-8	Sequence 8, Appl
45	179.5	3.0	1180	3 US-08-367-264-8	Sequence 8, Appl

ALIGNMENTS

RESULT 1
US-08-956-242-13
: Sequence 13, Application US/08956242C
: Patent No. 5986081
: GENERAL INFORMATION:
: APPLICANT: Gametxzy, Barry S.
: TITLE OF INVENTION: Polynucleotides Encoding Herg-3
: FILE REFERENCE: 960296.94550
: CURRENT APPLICATION NUMBER: US/08/956,242C
: NUMBER OF SEQ ID NOS: 13
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 13
: LENGTH: 1159
: TYPE: PRT
: ORGANISM: Homo sapien
: US-08-956-242-13

Query Match	100.0%;	Score 6079;	DB 2;	Length 1159;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 1159;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	MPVRGHVAPQNTFLDTIIRKFEQSRKFIITANRVCNAVITYCNDGFCELGYSRAEVM	60	
DB	1	MPVRGHVAPQNTFLDTIIRKFEQSRKFIITANRVCNAVITYCNDGFCELGYSRAEVM	60	
QY	61	ORPCTDPLHGPRTORRAAOIAQALGAERKVEIAFYRKDSCFLCTVNVVKNEDG	120	
DB	61	ORPCTDPLHGPRTORRAAOIAQALGAERKVEIAFYRKDSCFLCTVNVVKNEDG	120	
QY	121	AVIFILINFEVYMEKDWGSPAHDTNHRGPTSWLABGRATFPLKLPALATLARESSV	180	
DB	121	AVIFILINFEVYMEKDWGSPAHDTNHRGPTSWLABGRATFPLKLPALATLARESSV	180	
QY	181	RSGGAGAGAGAVVVDLTPAAPSSSESLALDEVYAMDNHVAGLGAERFALVPGSP	240	
DB	181	RSGGAGAGAGAVVVDLTPAAPSSSESLALDEVYAMDNHVAGLGAERFALVPGSP	240	
QY	241	PRSAAGLPSPRAHSLNDASGSSCLARTSRSCASVRRASSADIEARACVLPFP	300	
DB	241	PRSAAGLPSPRAHSLNDASGSSCLARTSRSCASVRRASSADIEARACVLPFP	300	
QY	301	RHASTGAMHPLRSGLNSTSDSDLVRYRTISKIPQITLNFVDLKGDFLASPTSDREITIA	360	
DB	301	RHASTGAMHPLRSGLNSTSDSDLVRYRTISKIPQITLNFVDLKGDFLASPTSDREITIA	360	
QY	361	PKIERHNHYTEKTOVLSIGADVLPEYKLOAPRIHRTILHYSFPFAVMDWMLLLVLY	420	
DB	361	PKIERHNHYTEKTOVLSIGADVLPEYKLOAPRIHRTILHYSFPFAVMDWMLLLVLY	420	

```

QY 421 TAVFTPYSAFLTKETEEGPATECGYACQPLAVVDLVDMFIVDILINFRTTYVNAME 480
DB 421 TAVFTPYSAFLTKETEEGPATECGYACQPLAVVDLVDMFIVDILINFRTTYVNAME 480
QY 481 EVVSHPGRIAVHYFEGWFLIDMVAIIPDILLIFSGSSEELIGLLTARLLRVYARKID 540
DB 481 EVVSHPGRIAVHYFEGWFLIDMVAIIPDILLIFSGSSEELIGLLTARLLRVYARKID 540
QY 541 RYSEGAVALFLMCTFALIAHMLACIWAIGNMEOPHMDSRIGMLHNLGDOIGKPYNS 600
DB 541 RYSEGAVALFLMCTFALIAHMLACIWAIGNMEOPHMDSRIGMLHNLGDOIGKPYNS 600
QY 601 GIGGSIIDKVTYALYFTFSSLTSGVGNVSPNTSEKIFISICVMLIGSLMAYSIFGVNS 660
DB 601 GIGGSIIDKVTYALYFTFSSLTSGVGNVSPNTSEKIFISICVMLIGSLMAYSIFGVNS 660
QY 661 AIIORLYSGTARYHTQMLRVREFIRFHQIIPNLRORLEBYRQHANSYNGIDMNAVLGKF 720
DB 661 AIIORLYSGTARYHTQMLRVREFIRFHQIIPNLRORLEBYRQHANSYNGIDMNAVLGKF 720
QY 721 PECLQADICLHNLNSLLOHCKPFRGATGCLRALAMKFKTTTHAPPGDTLVHAGDILLTALY 780
DB 721 PECLQADICLHNLNSLLOHCKPFRGATGCLRALAMKFKTTTHAPPGDTLVHAGDILLTALY 780
QY 781 FISRSISIELRGDVAVALIGKNDIFGEPLNLYARPGKSGNDYRALTCDLHKIHRDDLE 840
DB 781 FISRSISIELRGDVAVALIGKNDIFGEPLNLYARPGKSGNDYRALTCDLHKIHRDDLE 840
QY 841 VDMYPERSDHFWSSLEITFNLRODTNMTPGSPGSTELEGGFSROKRRKLSFRRRTDKOTE 900
DB 841 VDMYPERSDHFWSSLEITFNLRODTNMTPGSPGSTELEGGFSROKRRKLSFRRRTDKOTE 900
QY 901 QPGEVSALGPRAGAGPSSRGPRGPGWGESPPSSGSPSSSEDEBPGRSSSPLRLVPPSS 960
DB 901 QPGEVSALGPRAGAGPSSRGPRGPGWGESPPSSGSPSSSEDEBPGRSSSPLRLVPPSS 960
QY 961 PRPGEPPGGEPLMDCEKSSDTCNPLSGAFSGVSNITFSFGWDSRGROYQELPRCPAPTP 1020
DB 961 PRPGEPPGGEPLMDCEKSSDTCNPLSGAFSGVSNITFSFGWDSRGROYQELPRCPAPTP 1020
QY 1021 SLNLTPLSSPRRPGVDVESRLDALOROLNLETPLSDMAVLOLLOROMTLVYPAYSA 1080
DB 1021 SLNLTPLSSPRRPGVDVESRLDALOROLNLETPLSDMAVLOLLOROMTLVYPAYSA 1080
QY 1081 VTTPEPGPTSTSPLLPVSPPLTLTLDLSQVSOFMACELELPGABELPQEGFTTRLSPG 1140
DB 1081 VTTPEPGPTSTSPLLPVSPPLTLTLDLSQVSOFMACELELPGABELPQEGFTTRLSPG 1140
QY 1141 QLGALTSQPLHRHSGDPGS 1159
DB 1141 QLGALTSQPLHRHSGDPGS 1159

```

```

US-09-351-215-13
Query Match      100.0%; Score 6079; DB 3; Length 1159;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1159; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MPRRGHVAPONTFLDTIIRKREGOSRKFIINANAVENCAYICNDGFCGLGYRAEYM 60
DB 1 MPRRGHVAPONTFLDTIIRKREGOSRKFIINANAVENCAYICNDGFCGLGYRAEYM 60
QY 61 ORPCTDFLHGPRTRORRAAOIAQALLGAEKKVEIAFRKQSGFCCLVDVVPVKNDEG 120
DB 61 ORPCTDFLHGPRTRORRAAOIAQALLGAEKKVEIAFRKQSGFCCLVDVVPVKNDEG 120
QY 121 AVIMEILNEVYMEKDMGSPADPNHRGPPSWLAPRAKTFRLKLPALLALTARESSV 180
DB 121 AVIMEILNEVYMEKDMGSPADPNHRGPPSWLAPRAKTFRLKLPALLALTARESSV 180
QY 181 NSGGAGAGAPCAVYVDVLTTPAAPSSESLADEVTAMDNDHVAIGLPAEERRALVPGSP 240
DB 181 NSGGAGAGAPCAVYVDVLTTPAAPSSESLADEVTAMDNDHVAIGLPAEERRALVPGSP 240
QY 241 PSRAGOLPSPRAHSLNDPASGSCSIARTRESRESCASVRRASSADDIEMARAGVLPPEP 300
DB 241 PSRAGOLPSPRAHSLNDPASGSCSIARTRESRESCASVRRASSADDIEMARAGVLPPEP 300
QY 301 RHASTGAMHPLRSGLLNSTSDLVRYRTISKIPIQITLNFVDLKGDPFLASPTSDREITA 360
DB 301 RHASTGAMHPLRSGLLNSTSDLVRYRTISKIPIQITLNFVDLKGDPFLASPTSDREITA 360
QY 361 PKIKERTHNVTEKYOVLISLADVLEPKQAPRIHRRTILHAYSPKAWMDLILLLIY 420
DB 361 PKIKERTHNVTEKYOVLISLADVLEPKQAPRIHRRTILHAYSPKAWMDLILLLIY 420
QY 421 TAVFTPYSAFLTKETEEGPATECGYACQPLAVVDLVDMFIVDILINFRTTYVNAME 480
DB 421 TAVFTPYSAFLTKETEEGPATECGYACQPLAVVDLVDMFIVDILINFRTTYVNAME 480
QY 481 EVVSHPGRIAVHYFEGWFLIDMVAIIPDILLIFSGSSEELIGLLTARLLRVYARKID 540
DB 481 EVVSHPGRIAVHYFEGWFLIDMVAIIPDILLIFSGSSEELIGLLTARLLRVYARKID 540
QY 541 RYSEGAVALFLMCTFALIAHMLACIWAIGNMEOPHMDSRIGMLHNLGDOIGKPYNS 600
DB 541 RYSEGAVALFLMCTFALIAHMLACIWAIGNMEOPHMDSRIGMLHNLGDOIGKPYNS 600
QY 601 GIGGSIIDKVTYALYFTFSSLTSGVGNVSPNTSEKIFISICVMLIGSLMAYSIFGVNS 660
DB 601 GIGGSIIDKVTYALYFTFSSLTSGVGNVSPNTSEKIFISICVMLIGSLMAYSIFGVNS 660
QY 661 AIIORLYSGTARYHTQMLRVREFIRFHQIIPNLRORLEBYRQHANSYNGIDMNAVLGKF 720
DB 661 AIIORLYSGTARYHTQMLRVREFIRFHQIIPNLRORLEBYRQHANSYNGIDMNAVLGKF 720
QY 721 PECLQADICLHNLNSLLOHCKPFRGATGCLRALAMKFKTTTHAPPGDTLVHAGDILLTALY 780
DB 721 PECLQADICLHNLNSLLOHCKPFRGATGCLRALAMKFKTTTHAPPGDTLVHAGDILLTALY 780
QY 781 FISRSISIELRGDVAVALIGKNDIFGEPLNLYARPGKSGNDYRALTCDLHKIHRDDLE 840
DB 781 FISRSISIELRGDVAVALIGKNDIFGEPLNLYARPGKSGNDYRALTCDLHKIHRDDLE 840
QY 841 VDMYPERSDHFWSSLEITFNLRODTNMTPGSPGSTELEGGFSROKRRKLSFRRRTDKOTE 900
DB 841 VDMYPERSDHFWSSLEITFNLRODTNMTPGSPGSTELEGGFSROKRRKLSFRRRTDKOTE 900
QY 901 QPGEVSALGPRAGAGPSSRGPRGPGWGESPPSSGSPSSSEDEBPGRSSSPLRLVPPSS 960
DB 901 QPGEVSALGPRAGAGPSSRGPRGPGWGESPPSSGSPSSSEDEBPGRSSSPLRLVPPSS 960
QY 961 PRPGEPPGGEPLMDCEKSSDTCNPLSGAFSGVSNITFSFGWDSRGROYQELPRCPAPTP 1020
DB 961 PRPGEPPGGEPLMDCEKSSDTCNPLSGAFSGVSNITFSFGWDSRGROYQELPRCPAPTP 1020

```

```

RESULT 2
US-09-351-215-13
; Sequence 13, Application US/09351215
; Patent No. 6087488
; GENERAL INFORMATION:
; APPLICANT: Ganetzky, Barry S.
; TITLE OF INVENTION: Polynucleotides Encoding Herg-3
; FILE REFERENCE: 960296.94550
; CURRENT APPLICATION NUMBER: US/09/351,215
; CURRENT FILING DATE: 1999-07-12
; EARLIER APPLICATION NUMBER: 08/956,242
; EARLIER FILING DATE: 1997-10-22
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 13
; LENGTH: 1159
; TYPE: PRT
; ORGANISM: Homo sapien

```

QY	1021	SLNIPILSSPPRRRRRGVEESLDALQRLNLEFLRLSADATVYQLLRQMTLVPAYSA	1060
DB	1021		
QY	1081	VTTRGPGTSTSPLLPVSPLTLTLDLSISVSOPMACEELPGAGBELPQEGPTRLSLPG	1140
DB	1081		
QY	1141	QLGALTSQPLHRHGSDEGS	1159
DB	1141		
		1141 QLGALTSQPLHRHGSDEGS	1159
RESULT 3			
		US-09-226-012-2	
		; Sequence 2, Application us/092226012	
		; Patent No. 6207383	
		; GENERAL INFORMATION:	
		; APPLICANT: Keating, Mark T.	
		; APPLICANT: Spliwski, Igor	
		; TITLE OF INVENTION: MUTATIONS IN AND GENOMIC STRUCTURE OF HERG - A LONG QT	
		; TITLE OF INVENTION: SYNDROME GENE	
		; FILE REFERENCE: 2323-136	
		; CURRENT APPLICATION NUMBER: US/09/226,012	
		; CURRENT FILING DATE: 1999-01-06	
		; EARLIER APPLICATION NUMBER: 09/122,847	
		; EARLIER FILING DATE: 1998-07-27	
		; NUMBER OF SEQ ID NOS: 116	
		; SOFTWARE: PatentIn Ver. 2.0	
		; SEQ ID NO 2	
		; LENGTH: 1159	
		; TYPE: PRN	
		; ORGANISM: Homo sapiens	
		US-09-226-012-2	
		Query Match	
		Best Local Similarity 100.0%; Score 6079; DB 3; Length 1159;	
		Matches 1159; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY	1	MPVRGHAAPONTFLDTIIRKEGOSRKFIITANARVENCAYIYNDGFCELGYSRAEVM	60
DB	1	MPVRGHAAPONTFLDTIIRKEGOSRKFIITANARVENCAYIYNDGFCELGYSRAEVM	60
QY	61	QRPTCDPLHPRPQRRRAADIAALLGAERKVEIAFYRKDSCFCILDVYVYKNEG	120
DB	61	QRPTCDPLHPRPQRRRAADIAALLGAERKVEIAFYRKDSCFCILDVYVYKNEG	120
QY	121	AVIMFILFEVYMEKDMGSPAHNPTNHRGPTSMAPRAKTFPRKLPLALLALARESSV	180
DB	121	AVIMFILFEVYMEKDMGSPAHNPTNHRGPTSMAPRAKTFPRKLPLALLALARESSV	180
QY	181	RSGGAGAGABAVVVDLTPAAPSSSEIALDEVYAMDNHVAGLGAERRALVGPSP	240
DB	181	RSGGAGAGABAVVVDLTPAAPSSSEIALDEVYAMDNHVAGLGAERRALVGPSP	240
QY	241	PRSAAGOLPSRAHSLINDASGSSCSLARTSRSCASVRRASSADTLEARAGVLPSP	300
DB	241	PRSAAGOLPSRAHSLINDASGSSCSLARTSRSCASVRRASSADTLEARAGVLPSP	300
QY	301	RHASTGAMHPLRSGILNSTSDSDLVRYRTISKIPQITLNFVDLKGDPFLASPTSDREIIA	360
DB	301	RHASTGAMHPLRSGILNSTSDSDLVRYRTISKIPQITLNFVDLKGDPFLASPTSDREIIA	360
QY	361	PKIKERTINTEKYTOVLSLCAVDLPREKIDAPRIHRWTILHYSFKAVMWMLILLVIY	420
DB	361	PKIKERTINTEKYTOVLSLCAVDLPREKIDAPRIHRWTILHYSFKAVMWMLILLVIY	420
QY	421	TAVFPYSAAFLLKTEEGPRATEGVCOPLAVVDLVIMFIVDILINRTTYVANE	480
DB	421	TAVFPYSAAFLLKTEEGPRATEGVCOPLAVVDLVIMFIVDILINRTTYVANE	480
QY	481	EVSHSPGRIAVHYFKGWFILDMVAIPEDLLIFGSGSEELIGLKTARLLVRYARCLD	540

Db	Query Match	100.0%	Score 6079	DB 3	Length 1159
Db	Best Local Similarity	100.0%	Pred. No. 0		
Matches 1159	Conservative	0	Mismatches	0	Indels
				0	Gaps
				0	

QY 1 MPVRGHAAPONTFTDIIIRKFEKESRKFIIMANAVENCAYIYNDGFCCLGYSRAEYM 60
DB 1 MPVRGHAAPONTFTDIIIRKFEKESRKFIIMANAVENCAYIYNDGFCCLGYSRAEYM 60
QY 61 ORPCTCDLHGPRTORRAAOITAOALLGAEBKVEIAFYRKDSCFCLVDVVPYKNEBG 120
DB 61 ORPCTCDLHGPRTORRAAOITAOALLGAEBKVEIAFYRKDSCFCLVDVVPYKNEBG 120
QY 121 AVIMEFLNEEYMEKDMGSPAHDTNHRGPTSMILAPGAKTFRLKPLALLALTARESSV 180
DB 121 AVIMEFLNEEYMEKDMGSPAHDTNHRGPTSMILAPGAKTFRLKPLALLALTARESSV 180
QY 181 RSGGAGAGAPAVVVDLTPAASSESLADEVTAMDNHVAGLGPAREERALTVPKSP 240
DB 181 RSGGAGAGAPAVVVDLTPAASSESLADEVTAMDNHVAGLGPAREERALTVPKSP 240
QY 241 PSRAGOLPSPRAHSLNDPASGSSCLATRSRESCASVRRASSADDIEAMAGVLPKPP 300
DB 241 PSRAGOLPSPRAHSLNDPASGSSCLATRSRESCASVRRASSADDIEAMAGVLPKPP 300
QY 301 RHAATGAMHPLRSGLNSTSDLVRYRTISKIPQITLNFVDLKGDPFLASPTSDREIA 360
DB 301 RHAATGAMHPLRSGLNSTSDLVRYRTISKIPQITLNFVDLKGDPFLASPTSDREIA 360
QY 361 PRIKERTHNVEKTVQVLSLADVLPEYKLOAPRIHRMTILHYSFKAWMDLILLVY 420
DB 361 PRIKERTHNVEKTVQVLSLADVLPEYKLOAPRIHRMTILHYSFKAWMDLILLVY 420
QY 421 TAVFTPYSAFLKTEEBGPATTEGYACQPLAVVDLIVDIMEYDILINFRTTYVNAME 480
DB 421 TAVFTPYSAFLKTEEBGPATTEGYACQPLAVVDLIVDIMEYDILINFRTTYVNAME 480
QY 481 EVVSHPRGIAVHYFKGWLIDMVAIIPEDLIFGSGSEELIGLKTARLLRYVARKLD 540
DB 481 EVVSHPRGIAVHYFKGWLIDMVAIIPEDLIFGSGSEELIGLKTARLLRYVARKLD 540
QY 541 RYSEGAVALFLMCTFALIAHMLACIWAIGNMEOPHMDSRIGMLHNLGDOIGRPYNS 600
DB 541 RYSEGAVALFLMCTFALIAHMLACIWAIGNMEOPHMDSRIGMLHNLGDOIGRPYNS 600
QY 601 GUGGSIKDKYVTAIFYFTSSLSYGFVGNVSPNTSEKIFESICVMLIGSLMAYIFGNYS 660
DB 601 GUGGSIKDKYVTAIFYFTSSLSYGFVGNVSPNTSEKIFESICVMLIGSLMAYIFGNYS 660
QY 661 AIIOBLYSGTARYHTQMLRVREIFRHOIPNPLRORLEBYFOHANSYNGIDMNAVLKGF 720
DB 661 AIIOBLYSGTARYHTQMLRVREIFRHOIPNPLRORLEBYFOHANSYNGIDMNAVLKGF 720
QY 721 PECLADICLHNLRSLLQHKRFKATKGLRALMKFRTTHAPPGDITLVHAGDILLTLY 780
DB 721 PECLADICLHNLRSLLQHKRFKATKGLRALMKFRTTHAPPGDITLVHAGDILLTLY 780
QY 781 FLSRGSITELRGDVVAIILGKNDIFGEPLNTIYARPGKSGDVRALTYCOLHITHDDLE 840
DB 781 FLSRGSITELRGDVVAIILGKNDIFGEPLNTIYARPGKSGDVRALTYCOLHITHDDLE 840
QY 841 VLDMPPEFSDFHWSLEITFNLDNTMIPGSPSTELBEGFSROKRRKLSFFRRDCKDE 900
DB 841 VLDMPPEFSDFHWSLEITFNLDNTMIPGSPSTELBEGFSROKRRKLSFFRRDCKDE 900
QY 901 QPGEVSALGPRAGAGPSSRGPRGPGWGESPPSSGSPSSDEBGRSSPLRLVPSS 960
DB 901 QPGEVSALGPRAGAGPSSRGPRGPGWGESPPSSGSPSSDEBGRSSPLRLVPSS 960
QY 961 PRPPEPGEPLMDCEKSDTCNPLSGAFSGVSNITSEWGDSDGROQELPRCPAPRP 1020
DB 961 PRPPEPGEPLMDCEKSDTCNPLSGAFSGVSNITSEWGDSDGROQELPRCPAPRP 1020
QY 1021 SILANTPLSSPRGRPGDVESRLDALOROLNLETRLSADMATVILLQOMTLVPPAYSA 1080
DB 1021 SILANTPLSSPRGRPGDVESRLDALOROLNLETRLSADMATVILLQOMTLVPPAYSA 1080

QY 1081 VTBPGPPTSTPLPVSPPLPTLTDLSQVSQFACBCELPBGAPBLPQEGPTRLSPG 1140
DB 1081 VTBPGPPTSTPLPVSPPLPTLTDLSQVSQFACBCELPBGAPBLPQEGPTRLSPG 1140
QY 1141 OLGALTSOPLHRHSGDPS 1159
DB 1141 OLGALTSOPLHRHSGDPS 1159
RESULT 5
US-09-358-383C-10
; Sequence 10, Application US/09358383C
; Patent No. 6518398
; GENERAL INFORMATION:
; APPLICANT: Curtiss, Roy A.J.
; TITLE OF INVENTION: NOVEL POTASSIUM CHANNEL MOLECULES AND USES THEREFOR
; FILE REFERENCE: MNI-055CP
; CURRENT APPLICATION NUMBER: US/09/358, 383C
; PRIOR FILING DATE: 1999-07-21
; PRIOR FILING DATE: 1998-07-21
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 10
; LENGTH: 1159
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-358-383C-10
Query Match 100.0%; Score 6079; DB 4; Length 1159;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1159; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MPVRGHAAPONTFTDIIIRKFEKESRKFIIMANAVENCAYIYNDGFCCLGYSRAEYM 60
DB 1 MPVRGHAAPONTFTDIIIRKFEKESRKFIIMANAVENCAYIYNDGFCCLGYSRAEYM 60
QY 61 ORPCTCDLHGPRTORRAAOITAOALLGAEBKVEIAFYRKDSCFCLVDVVPYKNEBG 120
DB 61 ORPCTCDLHGPRTORRAAOITAOALLGAEBKVEIAFYRKDSCFCLVDVVPYKNEBG 120
QY 121 AVIMEFLNEEYMEKDMGSPAHDTNHRGPTSMILAPGAKTFRLKPLALLALTARESSV 180
DB 121 AVIMEFLNEEYMEKDMGSPAHDTNHRGPTSMILAPGAKTFRLKPLALLALTARESSV 180
QY 181 RSGGAGAGAPAVVVDLTPAASSESLADEVTAMDNHVAGLGPAREERALTVPKSP 240
DB 181 RSGGAGAGAPAVVVDLTPAASSESLADEVTAMDNHVAGLGPAREERALTVPKSP 240
QY 241 PSRAGOLPSPRAHSLNDPASGSSCLATRSRESCASVRRASSADDIEAMAGVLPKPP 300
DB 241 PSRAGOLPSPRAHSLNDPASGSSCLATRSRESCASVRRASSADDIEAMAGVLPKPP 300
QY 301 RHAATGAMHPLRSGLNSTSDLVRYRTISKIPQITLNFVDLKGDPFLASPTSDREIA 360
DB 301 RHAATGAMHPLRSGLNSTSDLVRYRTISKIPQITLNFVDLKGDPFLASPTSDREIA 360
QY 361 PRIKERTHNVEKTVQVLSLADVLPEYKLOAPRIHRMTILHYSFKAWMDLILLVY 420
DB 361 PRIKERTHNVEKTVQVLSLADVLPEYKLOAPRIHRMTILHYSFKAWMDLILLVY 420
QY 421 TAVFTPYSAFLKTEEBGPATTEGYACQPLAVVDLIVDIMEYDILINFRTTYVNAME 480
DB 421 TAVFTPYSAFLKTEEBGPATTEGYACQPLAVVDLIVDIMEYDILINFRTTYVNAME 480
QY 481 EVVSHPRGIAVHYFKGWLIDMVAIIPEDLIFGSGSEELIGLKTARLLRYVARKLD 540
DB 481 EVVSHPRGIAVHYFKGWLIDMVAIIPEDLIFGSGSEELIGLKTARLLRYVARKLD 540
QY 541 RYSEGAVALFLMCTFALIAHMLACIWAIGNMEOPHMDSRIGMLHNLGDOIGRPYNS 600
DB 541 RYSEGAVALFLMCTFALIAHMLACIWAIGNMEOPHMDSRIGMLHNLGDOIGRPYNS 600

Qy	601	GLGPESTKDKVTALYFTFSSLTSGFNVSPNTNSEKIFSCVWLISGLVASTFEGNVS	660	OTHER INFORMATION: Unidentified at time of filing
Db	601	GLGPESTKDKVTALYFTFSSLTSGFNVSPNTNSEKIFSCVWLISGLVASTFEGNVS	660	FEATURE: NAME/KEY: unsure
Qy	661	AIIOXLVSSTARHTOMLRVREFIRFHQIOPNPLROLEEYFOHAMSYNGIDMANVLGCF	720	OTHER INFORMATION: Unidentified at time of filing
Db	661	AIIOXLVSSTARHTOMLRVREFIRFHQIOPNPLROLEEYFOHAMSYNGIDMANVLGCF	720	FEATURE: NAME/KEY: unsure
Qy	721	PECLQADICHLNLSLOHCKPFRGATKGCRLALAMKFKTHAPPGDTLVHAGDILLTALY	780	OTHER INFORMATION: Unidentified at time of filing
Db	721	PECLQADICHLNLSLOHCKPFRGATKGCRLALAMKFKTHAPPGDTLVHAGDILLTALY	780	FEATURE: NAME/KEY: unsure
Qy	781	FISGSIETILGDDVVALTGKNDIFGEPLNLYARPGKNGDVRAITYCDLKHHRDDLLE	840	OTHER INFORMATION: Unidentified at time of filing
Db	781	FISGSIETILGDDVVALTGKNDIFGEPLNLYARPGKNGDVRAITYCDLKHHRDDLLE	840	FEATURE: NAME/KEY: unsure
Qy	841	VLDWYPEFSDFHWSLEITFNLRDNTMIPGSPSTELGCFSRQKRLTSFRRTDKDTE	900	OTHER INFORMATION: Unidentified at time of filing
Db	841	VLDWYPEFSDFHWSLEITFNLRDNTMIPGSPSTELGCFSRQKRLTSFRRTDKDTE	900	FEATURE: NAME/KEY: unsure
Qy	901	QPGVYSALGPGRAGAPSSRGRPGGPGWGESPPSGSPSESEDECGRSSPLRLVPSS	960	OTHER INFORMATION: Unidentified at time of filing
Db	901	QPGVYSALGPGRAGAPSSRGRPGGPGWGESPPSGSPSESEDECGRSSPLRLVPSS	960	FEATURE: NAME/KEY: unsure
Qy	961	PRPGEPPGSGPLWEDCKSSDTCNPLSGAGSVSNITFSFGDSRGROYELPRCPAPTP	1020	OTHER INFORMATION: Unidentified at time of filing
Db	961	PRPGEPPGSGPLWEDCKSSDTCNPLSGAGSVSNITFSFGDSRGROYELPRCPAPTP	1020	FEATURE: NAME/KEY: unsure
Qy	1021	SLNLIPLSSPPRRPGDVESRLDALQRLNLETRLSADMAVTLQLOROMTLVPAYSA	1080	OTHER INFORMATION: Unidentified at time of filing
Db	1021	SLNLIPLSSPPRRPGDVESRLDALQRLNLETRLSADMAVTLQLOROMTLVPAYSA	1080	FEATURE: NAME/KEY: unsure
Qy	1081	VTPGPGPTSTSPLLPVSPLEPTLTLDSLSQVSGFMACEBCLPPGAPELPQEGPTRLSLPG	1140	OTHER INFORMATION: Unidentified at time of filing
Db	1081	VTPGPGPTSTSPLLPVSPLEPTLTLDSLSQVSGFMACEBCLPPGAPELPQEGPTRLSLPG	1140	FEATURE: NAME/KEY: unsure
Qy	1141	QLGALTSGPLHRHGSDDPS	1159	OTHER INFORMATION: Unidentified at time of filing
Db	1141	QLGALTSGPLHRHGSDDPS	1159	FEATURE: NAME/KEY: unsure
RESULT 6				
US-08-956-242-4				
Sequence 4, Application US/08956242C				
Patent No. 5986081				
GENERAL INFORMATION:				
APPLICANT: Ganetzky, Barry S.				
APPLICANT: Titus, Steven A.				
TITLE OF INVENTION: Polynucleotides Encoding Herg-3				
FILE REFERENCE: 960296.94550				
CURRENT APPLICATION NUMBER: US/08/956.242C				
NUMBER OF SEQ ID NOS: 13				
SOFTWARE: PatentIn Ver. 2.0				
SEQ ID NO 4				
LENGTH: 888				
TYPE: PRT				
ORGANISM: Homo sapien				
FEATURE:				
NAME/KEY: unsure				
LOCATION: (133)				
OTHER INFORMATION: Unidentified at time of filing				
FEATURE:				
NAME/KEY: unsure				
LOCATION: (181)				
OTHER INFORMATION: Unidentified at time of filing				
FEATURE:				
NAME/KEY: unsure				
LOCATION: (438)				
OTHER INFORMATION: Unidentified at time of filing				
FEATURE:				
NAME/KEY: unsure				
LOCATION: (439)				

```
FEATURE:
NAME/KEY: unsure
LOCATION: (637)
OTHER INFORMATION: Unidentified at time of filing
FEATURE:
NAME/KEY: unsure
LOCATION: (646)
OTHER INFORMATION: Unidentified at time of filing
FEATURE:
NAME/KEY: unsure
LOCATION: (653)
OTHER INFORMATION: Unidentified at time of filing
FEATURE:
NAME/KEY: unsure
LOCATION: (656)
OTHER INFORMATION: Unidentified at time of filing
FEATURE:
NAME/KEY: unsure
LOCATION: (660)
OTHER INFORMATION: Unidentified at time of filing
FEATURE:
NAME/KEY: unsure
LOCATION: (661)
OTHER INFORMATION: Unidentified at time of filing
FEATURE:
NAME/KEY: unsure
LOCATION: (662)
OTHER INFORMATION: Unidentified at time of filing
US-08-956-242-4
```

Query Match 40.1%; Score 2436.5; DB: 2; Length 888;
Best Local Similarity 57.4%; Fred. No. 9.1e-197;
Matches 528; Conservative 71; Mismatches 166; Indels 155; Gaps 20;

```
QY 358 IIAKIKERTHNVEKYOVSLGADVPEYKLOAPRTHRTILHSPFKAWMDLILL 417
DB 1 IIAKIKERTHNVEKYOVSLGADVPEYKLOAPRTHRTILHSPFKAWMDLILL 60
QY 418 VIYRAVTPYSAFLKTEEGPATECGYACQPLAVVDLVDIMFIVDILINFTTYVN 477
DB 61 VIYRAVTPYSAFLKTEEGPATECGYACQPLAVVDLVDIMFIVDILINFTTYVN 119
QY 478 ANEEVSHRGIAVYEFKGFIDVVAIPDILLFGSGSEE--LIGLLTARLLRVR 534
DB 120 QNEEVSDPAKIAHYFGWFLIDVVAIPDILLFGSGSEETTLIGLLTARLLRVR 179
QY 535 VARKIDRYSEYGAUFLMCTFALIAHMLACIWIATNMOPHMSRIGMLHNDQIG 594
DB 180 VARKIDRYSEYGAUFLMCTFALIAHMLACIWIATNMOPHMSRIGMLHNDQIG 239
QY 595 KPYN-SSGLGSPSIRKDYVYALYFTSSLTSGVFGVNSPNTNSEKIFSICVWLIGSLMYA 653
DB 240 KRYNDSOSSSPSIRKDYVYALYFTSSLTSGVFGVNSPNTNSEKIFSICVWLIGSLMYA 299
QY 654 SIFGVSAIIORLYSGTARYTOMRVREFIRPHOIPMLRORLEEVFOHMSYNGIDM 713
DB 300 SIFGVSAIIORLYSGTARYTOMRVREFIRPHOIPMLRORLEEVFOHMSYNGIDM 359
QY 714 NAVLKGPEECLOADICHLNRSLLQHCPRFGATGCRALAMKFTTHAPGDVLVHAG 773
DB 360 NAVLKGPEECLOADICHLNRSLLQHCPRFGATGCRALAMKFTTHAPGDVLVHAG 419
QY 774 DLTLALYISGSLIIRGDDVVAALGKNDIFGEPLNLVYARPGSNGDVALTYCDLHKI 833
DB 420 DLTLALYISGSLIIRGDDVVAALGKNDIFGEPLNLVYARPGSNGDVALTYCDLHKI 479
QY 834 HRDLEVLDMYPERSDHFMSSLETFELRDTNM---IPSGPSLELGGSSROKRRKLS 890
DB 480 QREDLEVLDMYPERSDHFMSSLETFELRDTNM---IPSGPSLELGGSSROKRRKLS 539
QY 891 FRRRTDKD--TEOPGE-VSALGPGRAGAPSSRCRPGPGWESP-----SSGPPSP 938
DB 540 FESGGEKENSINDEPDSADTIRHVOSSXXPLKRRKXDPPLXSPPLMAMKXLFSLVDSX 599
```

```
QY 939 ESSDEGFGRSSPLRLVFPSSPRPGEPPGEPLMEDCEKSSDTCNPLSGAFSGVSNIF 998
DB 600 XIGEXASG-----LFXXTCPPPGE-----CPYIXXPPLAXXS--- 630
QY 999 SFWGDSRGRYOUELR-----CPA-----PPRSLINPL----- 1027
DB 631 --WLPYLGXK-NSLPRLXPLCPSSXGFLXXVPTSGRMHIDKRSKCKDITDMRWS 687
QY 1028 -----SPGRRRP-----GDVESRLDALORNLRLTSLADMAT 1062
DB 688 ERENANHPREDSPSALORAMGISETSDDLTYGEBQRDLLOBLNRLSOMTTDIOF 747
QY 1063 VLQLOQOMTLVPPASVATTPPGPSTSPLL-----PVSPLP--LTLDLSQVSOF 1114
DB 748 ILQLOQOTVVPAYSMVTA--GSEYORPILQIMRTSQPEASIKTDRSFPSSQCEPF 804
QY 1115 MACCELPPGAPLPOEG-----PVRRLSLP 1139
DB 805 LDLEKSKLSKSESLSGGVHLNTASEDNLTSLKODSDLSLEHLRQRTYVPIRHPSLP 864
QY 1140 -GOLGALTSOPLHRHSGDPG 1158
DB 865 DSSLSTVGIVGLHRHVSDDPG 884
```

```
RESULT 7
US-09-351-215-4
; Sequence 4, Application US/09351215
; Patent No. 6087488
; GENERAL INFORMATION:
; APPLICANT: Ganetzky, Barry S.
; TITLE OF INVENTION: Polynucleotides Encoding Herg-3
; FILE REFERENCE: 960296.94550
; CURRENT FILING DATE: 1999-07-12
; EARLIER APPLICATION NUMBER: 08/956,242
; EARLIER FILING DATE: 1997-10-22
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 4
; LENGTH: 888
; TYPE: PRT
; ORGANISM: Homo sapien
; NAME/KEY: unsure
; LOCATION: (133)
; OTHER INFORMATION: Unidentified at time of filing
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (181)
; OTHER INFORMATION: Unidentified at time of filing
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (438)
; OTHER INFORMATION: Unidentified at time of filing
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (439)
; OTHER INFORMATION: Unidentified at time of filing
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (567)
; OTHER INFORMATION: Unidentified at time of filing
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (568)
; OTHER INFORMATION: Unidentified at time of filing
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (571)
; OTHER INFORMATION: Unidentified at time of filing
```


FEATURE:
NAME/KEY: unsure
LOCATION: (575)
OTHER INFORMATION: Unidentified at time of filing
FEATURE:
NAME/KEY: unsure
LOCATION: (580)
OTHER INFORMATION: Unidentified at time of filing
FEATURE:
NAME/KEY: unsure
LOCATION: (589)
OTHER INFORMATION: Unidentified at time of filing
FEATURE:
NAME/KEY: unsure
LOCATION: (593)
OTHER INFORMATION: Unidentified at time of filing
FEATURE:
NAME/KEY: unsure
LOCATION: (599)
OTHER INFORMATION: Unidentified at time of filing
FEATURE:
NAME/KEY: unsure
LOCATION: (600)
OTHER INFORMATION: Unidentified at time of filing
FEATURE:
NAME/KEY: unsure
LOCATION: (608)
OTHER INFORMATION: Unidentified at time of filing
FEATURE:
NAME/KEY: unsure
LOCATION: (610)
OTHER INFORMATION: Unidentified at time of filing
FEATURE:
NAME/KEY: unsure
LOCATION: (611)
OTHER INFORMATION: Unidentified at time of filing
FEATURE:
NAME/KEY: unsure
LOCATION: (621)
OTHER INFORMATION: Unidentified at time of filing
FEATURE:
NAME/KEY: unsure
LOCATION: (624)
OTHER INFORMATION: Unidentified at time of filing
FEATURE:
NAME/KEY: unsure
LOCATION: (625)
OTHER INFORMATION: Unidentified at time of filing
FEATURE:
NAME/KEY: unsure
LOCATION: (628)
OTHER INFORMATION: Unidentified at time of filing
FEATURE:
NAME/KEY: unsure
LOCATION: (629)
OTHER INFORMATION: Unidentified at time of filing
FEATURE:
NAME/KEY: unsure
LOCATION: (634)
OTHER INFORMATION: Unidentified at time of filing
FEATURE:
NAME/KEY: unsure
LOCATION: (637)
OTHER INFORMATION: Unidentified at time of filing
FEATURE:
NAME/KEY: unsure
LOCATION: (646)
OTHER INFORMATION: Unidentified at time of filing
FEATURE:
NAME/KEY: unsure
LOCATION: (653)
OTHER INFORMATION: Unidentified at time of filing
FEATURE:

NAME/KEY: unsure
LOCATION: (656)
OTHER INFORMATION: Unidentified at time of filing
FEATURE:
NAME/KEY: unsure
LOCATION: (660)
OTHER INFORMATION: Unidentified at time of filing
FEATURE:
NAME/KEY: unsure
LOCATION: (661)
OTHER INFORMATION: Unidentified at time of filing
FEATURE:
NAME/KEY: unsure
LOCATION: (662)
OTHER INFORMATION: Unidentified at time of filing
US-09-351-215-4

Query Match 40.1%; Score 2436.5; DB 3; Length 888;
Best Local Similarity 57.4%; Pred. No. 9.1e-197;
Matches 528; Conservative 71; Mismatches 166; Indels 155; Gaps 20;

Qy 358 IIAPIKERTHNWTEKTVOLISGADVLPEXKLOAPRIHWTILHYSPPKAWMDLILL 417
Db 1 IIAPIKERTHNWTEKTVOLISGADVLPEXKLOAPRIHWTILHYSPPKAWMDLILL 60
Qy 418 VIYTAFTPTSAARLNDREE-QKRECGYSCSPLVNDLIVDLMFTLINFRTYYVN 477
Db 61 VIYTAFTPTSAARLNDREE-QKRECGYSCSPLVNDLIVDLMFTLINFRTYYVN 119
Qy 478 ANEYVSHPGRIAVHYKGFLLDMVAAPDILLFSGSSE--LGLKTAFLHLVR 534
Db 120 QNEEVSDPAKIAHYKGFLLDMVAAPDILLFSGSSE--LGLKTAFLHLVR 179
Qy 535 VARKLDYSEYGAFLVLMCTFALIAHMLACIYAIAGNMOPHMSRIGLHNGOIG 594
Db 180 VARKLDYSEYGAFLVLMCTFALIAHMLACIYAIAGNMOPHMSRIGLHNGOIG 239
Qy 595 KPVN-SSGLGPPSIRKDYVYALYFTFSSLSVGVGNVSPNTSEKIFSICVMLGSLMYA 653
Db 240 KRYNDSOSSSPSIRKDYVYALYFTFSSLSVGVGNVSPNTSEKIFSICVMLGSLMYA 299
Qy 654 SIFGNVSAIIORLVSIGARHTOMLRYREFRFOIENPLRQREYFOHMASTNGIDM 713
Db 300 SIFGNVSAIIORLVSIGARHTOMLRYREFRFOIENPLRQREYFOHMASTNGIDM 359
Qy 714 NAVLKGPECLADICHLNRSILOHCKPPRGATKGLRALAMFKTTTHAPGDTLVHAG 773
Db 360 NAVLKGPECLADICHLNRSILOHCKPPRGATKGLRALAMFKTTTHAPGDTLVHAG 419
Qy 774 DLTLALYISRGSIELRGDVVAIILGKNDIFGEPLNLVYAPGKSGNDVRLTYCDLHKI 833
Db 420 DLTLALYISRGSIELRGDVVAIILGKNDIFGEPLNLVYAPGKSGNDVRLTYCDLHKI 479
Qy 834 HRDDLVLVDYPPFSHFMSSEITFNLRDTNM---IPSPGSTLEEGESRORRKLK 890
Db 480 QREDLVLVDYPPFSHFMSSEITFNLRDTNM---IPSPGSTLEEGESRORRKLK 539
Qy 891 FRRRTDKD--TEOGE-VSALGPRAGSPSSRGPRGPMWESP-----SSGSSP 938
Db 540 FRRRTDKD--TEOGE-VSALGPRAGSPSSRGPRGPMWESP-----SSGSSP 599
Qy 939 ESSEDEGPRGSSPLRLVPESSPRPGEPGDEPLMDCEKSDTCNPLSGAFSGVSNIF 998
Db 600 XIGASG-----LXFXTCPPPE-----CPYKXPLAXXS--- 630
Qy 999 SFMGDSRGROYOLPR-----CPA-----TPPSLNTPL----- 1027
Db 631 --WLPXLGKX-NSLPRLEXPFLCPSXKXGLXXXVPTSGMHLDKRSHSGCKITDMRSW 687
Qy 1028 -----SSGGRPR-----GDVESRLDALORNLRLTSLADMAT 1062
Db 688 ERENAHQPEDSSSALQRAAMGISTESDLITTEVEORLDLLOEQLNRLESQMTTIQT 747


```

Db 138 -----VKTP-----EDKEDKVK-----151
QY 240 PPRSAQOLPPRAHSLNPDASGSSCSLARTSRSCASVYRASSADDIEMRAGVLP 299
152 -----GRSRACTHSDASARRRA-----169
QY 300 PRAHSTAGMPLRSGSLNSTSDLVRYRTISKIPQITLNFVDLKGDFLSPSDBREI 359
170 -----VLYHISGHQORREKNKLIKINNVEVDKP-----197
QY 360 APKIKERTHNTEKVTQVLSGADVLEPKLOAPRIHRMTLHSPKAVMDMLILLVI 419
198 -----APPEYKVSDAKSKSFILHSTFGAGDMWILLATF 233
QY 420 YTAVFTEYSAFLLKETEGRPATECGYACOPLAVDLIDIMFIVDILINFRTTYNAN 479
234 YVAATVYVYVNCFIENDLSTRTST-----TVSDIAVEILFDILINFRTTYYSKS 284
QY 480 EEVYSHFGRLAVYFKKMFILDMVAIIPEDLL-IFGSGSEELIGLKTARLLRVYARK 538
285 GQVTFEARSICIHVYTWTFIIDLIALPFDLLYAFNTVYVSLVHLKTVRLRLRLLOK 344
QY 539 LDRSEYGAAVFLMCTFALIAHMACIWAIGNMDOPH--MDSRIGWLHNLGDQIK 595
345 LDRISQSTIVLILMSNFALLAHMACIWTYIGKMEREDNSLKWEGWHLHEGKRLES 404
QY 596 P-YNNSGLGSPSIRKDYVYALYFTFSSITSVGFNVSPTNSEKIFISCVMLISLWYAS 654
405 PYYGNNTLGGPSISAVIAIALYFTLSLTSVGFNVSANTDAEKIFISCTMLIALMHAL 464
QY 655 IFGNVSAIIQRLYSGTARYTOMLRVREFIFHOIPNPLRORLEEYFOHANSYNGIDMN 714
465 VFGVNTAIIQMYSRMSLYHTRTKDLDFIRVHHLPOLKORMLEYEQTOTSVNNGIDSN 524
QY 715 AVLGFPECLOADICLHNSRLQHCKPRFGATGCTRALAMKFTTHAPRGDVLVHAGD 774
525 ELLKDPDELKSDITMHLNKEIILQ-LSLFECASRGCLSLHHTISFCAPAGEYLLROGD 583
QY 775 LITLALYISRSISILKGDVVAALIGKNDIFGEPLNLYARPGKNGVDRALTYDCLKIH 834
584 ALQAIYFVCSGSMELVLDMSVLAALIGKNDILGANLSIKDYIKTNADVKALTYDLOCII 643
QY 835 RDOLLEVDMYPERSDHFWSSL--EITFNLMD-----TNMIPGSPGSTELEGFSQOR 885
644 LKGFLEVGLDLYPEYAHNFVEDIQHDLTYNLRGESHSDVISRLSNKSNVSOSEPRGNGNIN 703
QY 886 KRKISFRRTDKDIOEGC--EVSALGPRAGAGPSSSRGPRG-----PWGE 929
704 KRLPISYE--DEEEEBEGEEBAVSLP-ICTRGSSSRNKVYSGNKAYLIGLSQOLASGT 760
QY 930 SPSSGP-----SSPESSEDEGPGRSSS-----PLRLVPPSPRPPGEPGPEPLMEDC 977
761 VPFSPISPIVRSKNSPKTKOEIDPNHNKRREKNLKLQSLTNMAGPDLSPRIYDGTEDG 820
QY 978 EKSSDTCNPLSGASGVSNITFSFGWDSRGROYELPRCPATPPLNIPLSSPGRRRGD 1037
821 NSSEB-----SQTFDF-GSEBRIR-----SEPRISIP---PLGD 848
QY 1038 VE-----SRDLAQOLNRLR-----TRLSDMATVLTQLOROGTLPVPA 1077
Db 849 PEIQAVALFIKAETKQOINKLNSEVITLQEVSQLKDMKNVIRLLENVLSPOOPS 905

```

```

; PRIOR APPLICATION NUMBER: USSN 09/119,855
; PRIOR FILING DATE: 1998-07-21
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 36
; LENGTH: 1102
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-09-358-383C-36

Query Match 25.5%; Score 1551; DB 4; Length 1102;
Best Local Similarity 32.6%; Pred. No. 8.3e-122;
Matches 383; Conservative 169; Mismatches 283; Indels 340; Gaps 30;

QY 1 MPVRGHVAPONTFLDTIIRKEGOSRKFITANARV-ENCAYIYCNDFCELCYSRAEV 59
1 MPVKGGLAPONTFLDTIATRFDGHSHFILANQVAKGPYIYVCSDFELAGFAREY 60
QY 60 MQRPTCDFLHGPRTORRAAQIQAALLGAERKVEYLAFTKKDSCFELCIVDVPVKNED 119
61 MOKSCSKFLFGEVETNQLQIEKSLSEKVEFGELIMFYKKNGAPFMCCLDIYPIRNEK 120
QY 120 GAVIMELINEFVMEKDWGSPADTNHRGPTSMALAPRAKTFRLKLPALLATARESS 179
121 GDVYVLFASF-----KDTITIKV-----138
QY 180 VRSGAGAGAPGAVVVDVLTTPAAPSSESLADEVTAMDNHVAAGLPAERRALVGPSS 239
139 -----KITSQK-----KEDR-----149
QY 240 PPRSAQOLPPRAHSLNPDASGSSCSLARTSRSCASVYRASSADDIEMRAGVLP 299
150 ---AKGR-----SRAGHPDSARRSR-----168
QY 300 PRAHSTAGMPLRSGSLNSTSDLVRYRTISKIPQITLNFVDLKGDFLSPSDBREI 359
169 -----AVLYHISGHQORREKNKLIKINNVEVDKP-----197
QY 360 APKIKERTHNTEKVTQVLSGADVLEPKLOAPRIHRMTLHSPKAVMDMLILLVI 419
198 -----APPEYKVSDAKSKSFILHSTFGAGDMWILLATF 233
QY 420 YTAVFTEYSAFLLKETEGRPATECGYACOPLAVDLIDIMFIVDILINFRTTYNAN 479
234 YVAATVYVYVNCFIENDLSTRTST-----TVSDIAVEILFDILINFRTTYYSKS 284
QY 480 EEVYSHFGRLAVYFKKMFILDMVAIIPEDLL-IFGSGSEELIGLKTARLLRVYARK 538
285 GQVTFEARSICIHVYTWTFIIDLIALPFDLLYAFNTVYVSLVHLKTVRLRLRLLOK 344
QY 539 LDRSEYGAAVFLMCTFALIAHMACIWAIGNMDOPH--MDSRIGWLHNLGDQIK 595
345 LDRISQSTIVLILMSNFALLAHMACIWTYIGKMEREDNSLKWEGWHLHEGKRLES 404
QY 596 P-YNNSGLGSPSIRKDYVYALYFTFSSITSVGFNVSPTNSEKIFISCVMLISLWYAS 654
405 PYYGNNTLGGPSISAVIAIALYFTLSLTSVGFNVSANTDAEKIFISCTMLIALMHAL 464
QY 655 IFGNVSAIIQRLYSGTARYTOMLRVREFIFHOIPNPLRORLEEYFOHANSYNGIDMN 714
465 VFGVNTAIIQMYSRMSLYHTRTKDLDFIRVHHLPOLKORMLEYEQTOTSVNNGIDSN 524
QY 715 AVLGFPECLOADICLHNSRLQHCKPRFGATGCTRALAMKFTTHAPRGDVLVHAGD 774
525 ELLKDPDELKSDITMHLNKEIILQ-LSLFECASRGCLSLHHTISFCAPAGEYLLROGD 583
QY 775 LITLALYISRSISILKGDVVAALIGKNDIFGEPLNLYARPGKNGVDRALTYDCLKIH 834
584 ALQAIYFVCSGSMELVLDMSVLAALIGKNDILGANLSIKDYIKTNADVKALTYDLOCII 643
QY 835 RDOLLEVDMYPERSDHFWSSL--EITFNLMD-----TNMIPGSPGSTELEGFSQOR 885
644 LKGFLEVGLDLYPEYAHNFVEDIQHDLTYNLRGESHSDVISRLSNKSNVSOSEPRGNGNIN 703

```

```

RESULT 12
US-09-358-383C-36
; Sequence 36, Application US/09358383C
; Patent No. 6518398
; GENERAL INFORMATION:
; APPLICANT: Curtis, Roy A.J.
; TITLE OF INVENTION: NOVEL POTASSIUM CHANNEL MOLECULES AND USES THEREFOR
; FILE REFERENCE: NMI-055CP
; CURRENT APPLICATION NUMBER: US/09/358,383C
; CURRENT FILING DATE: 1999-07-21

```

```

QY 881 !SRQRRKRLSRRTDXTDTEQPGVSLGPRACAGPSSRCRRPGGPMCESSSPSPSES 940
Db 700 GS-!KKRLPUSIVDEDEEEVEEETTSISP-----!YTRK-----SSVSH 738
QY 941 SEDEGPGRRSSPLRL-----VFSSP-----RPPGEP----- 968
Db 739 SKTGSSSKSYGLSLKQLTSGTVFPFHPISVSSANSPKTKQADAPPNGHTRKKNLQVL 798
QY 969 -----GGEPLM-----EDCEKSSDPCNLSAFSSVSNIFSFMDSSRRQVQELPRC 1015
Db 799 CSLCTAGAPPELSPRIVDIEDGNSSEET-----QTFPGSE-----QIRPEPRI 842
QY 1016 PAPPTSLNIPILSSPGRRRGDVE-----SRDLAQRLNRLE-----TRLSA 1058
Db 843 ---SPSL-----GESFICGAFLFKAEKTRQKQINKLNSSEVTLTQEVSGQSK 886
QY 1059 DMATVQLQLQKQMTLVPPPAISAVVTPPGGPISTSP 1093
Db 887 DMRSIMOLLENILSPQOPSOFC---SLHPSTIP 917

```

```

RESULT 13
US-09-343-494-1
; Sequence 1, Application US/09343494
; Patent No. 6413741
; GENERAL INFORMATION:
; APPLICANT: Jegla, Timothy J.
; APPLICANT: Wickenden, Alan
; APPLICANT: ICagen, Incorporated
; TITLE OR INVENTION: Human Elk, a Voltage-Gated Potassium Channel Subunit
; FILE REFERENCE: 0.18512-001320US
; CURRENT APPLICATION NUMBER: US/09/343,494
; EARLIER FILING DATE: 1999-06-30
; EARLIER APPLICATION NUMBER: US 60/091,469
; EARLIER FILING DATE: 1998-07-01
; EARLIER APPLICATION NUMBER: US 60/116,621
; EARLIER FILING DATE: 1999-01-21
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 1083
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: human Elk (hElk); Eag (eag) gene
; OTHER INFORMATION: potassium channel monomer protein
US-09-343-494-1

```

Query Match	25.4%	Score 1541.5	DB 4	Length 1083
Best Local Similarity	32.4%	Pred. No. 5.1e-121		
Matches	410	Conservative 143	Mismatches 330	Indels 377
				Gaps 35
QY	1	MPVRGHVAPONT	LDLTIRKKEGOSRKFTI	ANARVENC-AVIYCNDGFC
		1		
DB	1	MPARGLIAPONT	ELDTATRF	DGTHSNFVGNQVADLFVYVCCSDFC
QY	60	MORPCTGDLG	PTPTQRRAA	NOIVALLGAEE
DB	61	MORCAGCS	FLYGPDTSEL	YRQIKKALDEHKEF
QY	120	GAVIMFLIN	FEVMEKDNVGS	PAHDITNHRGPT
DB	121	GEVALFL	-----VSHKDI	-----SETKNRGSGP
QY	180	VRSGGAGAG	AGAVVVDVDTL	PAAPSESLALDE
DB	150	-----GG	-----	RRR
QY	240	PPRSAPGOL	PSPRAHSLN	LPDASGSSCLART
DB	155	-----	YGRARSKG	FNANRRRRSR

QY	360	RHAHSTGAMHPLRGLSLNSTDSDLVKRRITISKIPQITLINVVDLKGDPFLASPTSDREII	359
Db	185	PK-----GKHLNKGVEGEPN-----	201
QY	360	APKIKERTHNTYKTVQVLSIGADVLPEYKLOAPRIHRWTLIHXSPKAVMDLILDLVI	419
Db	202	-----LPEYKAAALIRKSPFILLHCGALRATPFGDILLATL	236
QY	420	YTAVFTPYSAFLKEEBGPATBEGYACOPPLAVDLVDIMFTVIDLINFRTTYNNAN	479
Db	237	YVAIVVPYSVCV-----STARPESA-----ARGPSPCDLAEVLEFILDIVLNFRTTFVSKS	288
QY	480	EEVYSHPRIVAHYKKGWFLIDMYAAIPEDL-----IFGSGSEELIGLTKTARILL	532
Db	289	GOVVFAPRPSICLHYVTYTMFLDVLDALEFDLILHAFKNVYFGA-----HLKTVRLRL	342
QY	533	VRVAKIDRYSEYGAVALLELMCFALIAHMLACIWAIGNMEOPHMS--RIGWMLNL	589
Db	343	LRLPLRIDRYSQYSAVVLITLMAVFPALLAHVACWVFYIGORELSESSLEPELIGWLOEL	402
QY	590	GDQIGKPY-----NSSG-----LGPSIKDKYITALYFPSS	621
Db	403	ARRLTTPYVLGRPRAGCNSSGOSDNCSSSSEANGTGLELLGGSLSRAYSATTSYFALSS	462
QY	622	LTSVGFVSPNTNSEKJFESICVWLIGSLMASTJFGVNASITQRYSGTARTYHOMLRVR	681
Db	463	LTSVGFVNASNTIDEKJFESICTMLIGALMAVYFGVNTALIQMYARFLYHSRTDLR	522
QY	682	BFIREHQIPNPLRQLEEYFQHAHNSYTXGIDMANVNLGFPICLOADICLHLNRSLLQHK	741
Db	523	DYIRHRIKPKLOKQMLEYFOATMAVNNGIDTTELLQSLPELRADIAHMLHKEVLO-LP	581
QY	742	PFKRGATKCLALAAKFKTTHAPRGDITLVHAGDILLTALYFSRSISIELRBDYVVALIG	801
Db	582	LFEASRCLALSTALRPACFTGGEYIHHOGDALQALFYVCSGSMELKGVLTALIGK	641
QY	802	NDIFGEPLNLRAPBCKNSGDYRALTYYCDLHKIHRDDLEVLDMYPEFSDHFWSSL-EIT	859
Db	642	GDILGCELPRRBOYKANADVKGLTYCLOQLAGLHDSLALYPAFAPRSGLRGELS	701
QY	860	FNLRDTNMTIPOSPOSTLEGGFSRORKKLSFRRTDKTBO-----	901
Db	702	YNL-----GAGGGEAEVDTSSLGDNLTMLTEBEKEPDGQGPVSPAPADEBSSPLLS	755
QY	902	PGEVSA-----LGGGRAGAP--SSRRRGPGMPGESSPSSPESEEDGPGRSSPL	953
Db	756	PGCTSSSSAAKLISRRAPRRPLRGGRKRPBAGALKAENAPSAP-----RALEGL	807
QY	954	RLVPESSRPGEPEPGEPE-LMEDCEKSSIDCNPLSGAFSGVSNISFPMWDSRROYOEL	1012
Db	808	RL-----PMPMNVNPPDLSPRVVDGIEDGCGSDQK-----FSP-----RVQSG	847
QY	1013	PRC-----PAFPSLINIPLSSPGRRPBGVDSR---LDALQOLNRLNLETRLSDMATV	1063
Db	848	PECSSSPSPGEBESGLTV-----PHGPSEARNTDILDKLRQAVELSLBOV-LQWREG	898
QY	1064	LOLOROMTIV-----	1082
Db	899	LOSLQOAVOLVLAHPRREGPCPRASGECPASTSGLOPLQCVYMGASSYCLQAPRAGSVLS	958
QY	1083	-----TPGPGPTST--SPLLPVSLP-----TLTUDLSQVSOFMACEELPQAP	1125
Db	959	GTWHPAPAGPPLIAMPWMPGPPASQSSWPMPATAFWMTSTSDSEPPASGDLCSEBSTPASP	1018
QY	1126	ELPQEG 1131	
Db	1019	PPSEEG 1024	

RESULT 14
US-09-600-776-2
; Sequence 2, Application US/09600776
; Patent No. 6326168

```

: GENERAL INFORMATION:
: APPLICANT: Yamouchi Pharmaceutical Co., Ltd.
: TITLE OF INVENTION: A novel potassium channel protein
: FILE REFERENCE: Y9903-PCT
: CURRENT APPLICATION NUMBER: US/09/600,776
: PRIOR FILING DATE: 1998-01-23
: PRIOR APPLICATION NUMBER: JP P1998-011434
: PRIOR FILING DATE: 1998-01-23
: PRIOR APPLICATION NUMBER: JP P1998-346198
: NUMBER OF SEQ ID NOS: 12
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 2
: LENGTH: 1083
: TYPE: PRT
: ORGANISM: Homo sapiens
US-09-600-776-2

Query Match      25.3%; Score 1540.5; DB 4; Length 1083;
Best Local Similarity 32.4%; Pred. No. 6,2e-121;
Matches 410; Conservative 143; Mismatches 336; Indels 377; Gaps 35;

QY      1 MPVRGHVAPONTFLDTIIRKFEQSRKFTIANARVENC-AVITYCNDGFCGLGYRAEV 59
      11 |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      1 MPAMRGILAPONTFLDTIATRFDTGTHSNFVIGNAQVAGLPVYVYCSGDFCLTGFSAEV 60

QY      60 MQRPTCDLHGPRTORRAAQAQALLGAERKVELAFYRKDGSCECLVDVVPVKNED 119
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      61 MQRCAQSFLYGPDTSFLVROQIRKALDEHKEFEKAEILLYRKSGLPFCMLLDVPIKNEK 120

QY      120 GAVIMFLNFEVMEKMGVSPAHDTNHRGPPTSMLAPGAKTRFLKLPALLALTARESS 179
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      121 GEVALFL-----VSHKDI-----SETKNRGGPDWKETG----- 149

QY      180 VRSGGAGAGACGAVVVDVLTTPAASSESLALDEVTAMDNHVAGLGAEEERALVGPSS 239
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      150 -----GG-----RRR----- 154

QY      240 PPSAPGOLPPSPRAHSLNPAGSSGSLARTSRSCASVRRASSADDIEAMRAGVLP 299
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      155 -----YGRARSKGFNANRRSRPA-----VLYHLSGHLQKQ 184

QY      300 PRAHSTGAMHPLRSGLNSTSDSLVYRITISKIPQTLTNVLDKGPFLASPTSDRETI 359
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      185 PK-----GKHKLNGVGEKPN----- 201

QY      360 APRIKERTHNVTEKVTQVLSIGADVLEPEYKLOAPRIHRWTILHSPKAVMDMLILLVI 419
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      202 -----LPEYKVAAIRKSPFILLHCGALRATWDGFIILATL 236

QY      420 YTAVFPTYSAAFLKETEGSPATECGYACQPLAVDLIVDIMEIVDLINFTTYVAN 479
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      237 YVAVTVYVYCV-----STARPSA-----ARGPSPVCDLAVEVLTILDIVLNFRTPYFSKS 288

QY      480 EEVYSHGRVAVHYFKGMFLIDMVAALPBDL-----IFGSSSEELIGLKARLR 532
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      289 GQVVPKSKTCLHVTWFLDVLDTALPBDLHAEKVNYFGA-----HLKTVRLRL 342

QY      533 VVARUKLDRTSEYGAVALFLMCTFALIAHMLACTIAYINMEQPHMS---RIGWHLN 589
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      343 LRLPLRDRYSQYSAVVLTLLMAVFALAHVACVWYIIORELESESESELPEIGWQEL 402

QY      590 GDOIGKPY-----NSSG-----LGSPSIRKDYVTALYTFSS 621
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      403 ARLETFPYLVYGRPPAGNSSGQSDNCSSEANGTGLLELGGPSLSAVITSLYFLASS 462

QY      622 LTVGRCNVPNPNNSEKIFISICVMLIGSLMAYASIFGVNSAIQRLYSGTARYHQMLRVR 681
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      463 LTVSGGVNSANVTDEKIFISICVMLIGSLMAYAVGVNTAIIQRMVARRPLYHSRTDRLR 522

QY      682 EEFIRFQIPLRLQRLLEEFQOHAMSYTNGIDMNAVLGKGFPECLOADCILLNRSLQCK 741
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      523 DYIRIRIRPKRLQRMLETFQATWAVNNGIDITTELLQSLPDELRLADIAHMLHKEVLDL 581
```

```

QY      742 PFRGATKGCIALAMKFKTTHAPPGDPLVHAGDLLTALYFISKSIETLRGDIVVAILGK 801
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      582 LFEAASRGCLALSLALRPACFTGEXYILHOGDALQALYFVCSGSMELKGVIALIIGK 641

QY      802 NDIFGEPLNLAYARKCKNSGVNDRALTYCDLKHIIHDDLLEVDMMPEPSDHWSSL--EIT 859
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      642 GDLIGCELPVRRQYKANADVKGLTYCVLQCLQLAGLHSDIALYPERAPRSRGLKRELS 701

QY      860 ENLRDTNMIPGSPSTELGEGFSRQRRKLSFRRTDKDTFO----- 901
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      702 YNL-----GAGGSAEYVDTSLSGDMTLMSTLEKETDEGQGTVPAPADESSPLS 755

QY      902 PGEVYA-----LGPGRAGACP--SSRGPRGPGWGESPPSGPSSPESEDEGPRSSSPL 953
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      756 PGCTSSSSAAKLLSPRTAPRPLRGCRGCRGACALKAAGSPAP-----RALEGL 807

QY      954 RLVPFSSPRRPGEPGEGP-LMEDEKSSDTCNPLSCAFSGVSNIFEFWMDSRQYQEL 1012
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      808 RL-----PPMWNVPDLSPRVVDIEEGCGSDQPK-----FSF-----RVGQSG 847

QY      1013 PRC-----PAPTSLNIPLSPPGRBRGDVSR-----LDALORQLRLLETRLSDAMATV 1063
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      848 PECSSSPSPGEGSLTY-----PHGFSRANTDTLDKROAVTELSQV-LQMRG 898

QY      1064 LQLLQROMTLY-----PPAYSAYT 1082
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      899 LQSLRAQVQVLAHREGPCPRASGEGPCPASTSGLLQPLCVDPGASSYCLQPPAGSVLS 958

QY      1083 -----TPGCGPTST-----SPLLVSPPL-----LTLDLSQVSGPMAFCELP 1125
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      959 GTWHPRRPGPPLMAPMPWGPASQSSPPWPAATATWSTSDSEPPASGDLCSEPTSP 1018

QY      1126 ELPOEG 1131
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      1019 PPSEEG 1024

RESULT 15
US-09-358-383C-2
Sequence 2, Application US/09358383C
Patent No. 6518398
: GENERAL INFORMATION:
: APPLICANT: Curtiss, Roy A.J.
: TITLE OF INVENTION: NOVEL POTASSIUM CHANNEL MOLECULES AND USES THEREFOR
: FILE REFERENCE: NMT-055CP
: CURRENT APPLICATION NUMBER: US/09/358,383C
: PRIOR FILING DATE: 1999-07-21
: PRIOR APPLICATION NUMBER: USSN 09/119,855
: NUMBER OF SEQ ID NOS: 36
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 2
: LENGTH: 1083
: TYPE: PRT
: ORGANISM: Macaca sp.
US-09-358-383C-2

Query Match      25.3%; Score 1540.5; DB 4; Length 1083;
Best Local Similarity 32.2%; Pred. No. 6,2e-121;
Matches 411; Conservative 146; Mismatches 344; Indels 377; Gaps 35;

QY      1 MPVRGHVAPONTFLDTIIRKFEQSRKFTIANARVENC-AVITYCNDGFCGLGYRAEV 59
      11 |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      1 MPAMRGILAPONTFLDTIATRFDTGTHSNFVIGNAQVAGLPVYVYCSGDFCLTGFSAEV 60

QY      60 MQRPTCDLHGPRTORRAAQAQALLGAERKVELAFYRKDGSCECLVDVVPVKNED 119
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      61 MQRCAQSFLYGPDTSFLVROQIRKALDEHKEFEKAEILLYRKSGLPFCMLLDVPIKNEK 120

QY      120 GAVIMFLNFEVMEKMGVSPAHDTNHRGPPTSMLAPGAKTRFLKLPALLALTARESS 179
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      121 GEVALFL-----VSHKDI-----SETKNRGGPDWK----- 145
```

```

QY 180 VRSAGAGAPGAVVVDVLTTPAAPSSESLALDEVAMDNDHAGLGAPEERRALVPGS 239
   : : : : :
Db 146 -KETSAG-----RRR-----154

QY 240 PPRSAPGOLPRAHSLNPDASGSSCLARTSRRESCASVRRASADDIEAMRAGVLPPP 299
   : : : : :
Db 155 -----YGRARSKGFNRRRSRA--VLYHLSGHLQK 184

QY 300 PRAHSTGAMHPLASGLNSTSDSLVRYRTISKIPQITLNFVLDKDPFLASPTSDREIT 359
   : : : : :
Db 185 PK-----GKHKLKGVFGEKPN-----201

QY 360 APRKERTHNTEKVTQVLSLADVLPEYKLOAPRIHRWTLHSPPKAVMDWLLLVY 419
   : : : : :
Db 202 -----LPEYVAAIRKSPFILLHCGALRATWDGFIILLATL 236

QY 420 YTAFTPYSAFLLKTEEGPPATECGYACQPLAVVDLIVDIMEIVDILINFRTYVNAV 479
   : : : : :
Db 237 YVAVTPYVCY---STARPSA-----ARGPPVCDLAVEVLFILDIVLNFRTFVSKS 288

QY 480 EEVYSHRGRIAVHYFKGFILDMVAIPDIL-----IFSGSEELIGLTKTARLLRL 532
   : : : : :
Db 289 GQVVFAPKSLCHYVTTWFLDVIYALPFDLHAFKVNYFEA-----HLKTVRLRL 342

QY 533 VVARKLDRYSEGAIVLFILMCTFALIAHMLACIWAYAIGNMOPHMDS---RIGWLNHL 589
   : : : : :
Db 343 LRLPLRLDRYSQYSAVYVTLMAVFPALIAHVACWFIYIQOREISESESELPEIGWLOEL 402

QY 590 GDQIGKPY-----NSSG-----LGGPSIKDKYVYALYFTSS 621
   : : : : :
Db 403 ARRLTEPYLVGRRRPAGNSSGQSDNCSSSEANGTGLLGGPSLRSAYITSLYFALSS 462

QY 622 LTVSGCNVSPNNTSEKIFISICVMLIGSLMYASIFGNVSAIIQRIYSGTARHTQMLRVR 681
   : : : : :
Db 463 LTVSGCNVSNANDTEKIFISICTMLIGALMAHVAVFGNVTALIQRYARRELYHSRTRDLR 522

QY 682 EFTIRHOIPNPLRQRLSEYFOHMAWSTYNGIDMNAVILKGPCELOADICLHLNRSILOHCK 741
   : : : : :
Db 523 DYIRIRIRKPLQRMLEFYQATWAVNNGIDITELLOSPLDELAMHLKEVLO-LP 581

QY 742 PFRGATKGLRALAMFKTTTHAPPGDTLVHAGDULTALYFISRGSIEILRGDVVAAILGK 801
   : : : : :
Db 582 LFEASRGLRALSLALRPAFCPTGEVLIHQDOLQALYFVCSGMEVLKGVYLAITLGK 641

QY 802 NDIFGEPLNLYARPKSNGDYRALTYCDLKHIRHDLLEYLDMPERSDHFWSL--EIT 859
   : : : : :
Db 642 GDILGCELPREQYVVKANADVKGITYCVLOQLAGLHDSLALYPEFAPRFSRGLRGELS 701

QY 860 FNLBDTMMIPGSPGSTELEGFSRQRRKLSFRRTDKDTEQ-----901
   : : : : :
Db 702 YNL-----GAGGSALEVDTSSLSGDNTLSTLEEKETDGEQPTVSPAPADESSPLLS 755

QY 902 PGEVSA-----LQPGRAGAP--SSRGRPGPGWGESPPSSPSESSEDEGPGRSSPL 953
   : : : : :
Db 756 PGCSSSSSAKALSPRTAPRPRIGRGRGRAGALAKAEGAPSAAP-----RALEGL 807

QY 954 RLYPFSSPPRPPGEGEP--LMEDEKSSPTCNPLSGAFSGVSNIFSFWDSDSRGYOEL 1012
   : : : : :
Db 808 RL-----PPMPWNVPPDLSPRVVDGIEDGCSQDPK-----FSF-----RMGQSG 847

QY 1013 PRC-----PAPTSLNIPILSSPGRRRRGVRESR---LDALQRLRLLETRLADMATV 1063
   : : : : :
Db 848 PECSSSPSPGESGLTV-----PHGSEARNTTLDLKLQAVMELSEQV--LQMRBG 898

QY 1064 LQTLQROMTLV-----PPAYSAYT 1082
   : : : : :
Db 899 LQSLRQAVOVLVLAHREGPCPRASGEGPCPASTSGLLOPLCYDTGASSYCLQPPRAGSVLS 958

QY 1083 -----TPGPGPTST-----SPLLVPSPPLTLTIDSLSQVQFMACEELPPGAPELPOEG 1131
   : : : : :
Db 959 GTWPHRRPBPPLMAPWPGCPRASQSSPMPRATAFMTS-----TSDSEPPASGDLCSGP 1012

```

```

QY 1132 PTRRLSLPGQLGALTSOP 1149
Db 1013 STPASPPSPSEGARTGPP 1030

```

Search completed: September 23, 2003, 17:41:04
Job time : 44 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 23, 2003, 17:37:05 ; Search time 455 Seconds
(without alignments)
379.269 Million cell updates/sec

Title: US-10-000-151B-3
Perfect score: 6079
Sequence: 1 MPVRGHVAPONTFLDTIIR.....COLGALTSQLRHGSDPGS 1159

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 556269 seqs, 148893369 residues

Total number of hits satisfying chosen parameters: 556269

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

Published Applications_AA.*
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/PCR_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/PCRUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep:*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*
18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	6079	100.0	1159	9	US-09-735-995-2
2	6079	100.0	1159	9	US-09-735-995-4
3	6079	100.0	1159	9	US-09-119-855-10
4	6079	100.0	1159	15	US-10-000-151B-3
5	6079	100.0	1159	15	US-10-193-692-5
6	6079	100.0	1159	15	US-10-185-867-10
7	6079	100.0	1159	15	US-10-174-613-49
8	3493.5	57.5	1177	15	US-10-193-692-4
9	3312.5	54.5	1195	15	US-10-192-440-10
10	3255	53.5	1196	15	US-10-255-532-2
11	3241	53.3	1186	15	US-10-193-692-2
12	3043.5	50.1	994	15	US-10-192-440-8
13	3001.5	49.4	958	12	US-10-323-430-21
14	3001.5	49.4	958	15	US-10-192-440-2
15	2982	49.1	950	15	US-10-192-440-9

16	1636	26.9	1017	10	US-09-965-830-6	Sequence 6, Appl1
17	1568	25.8	1107	15	US-10-185-867-16	Sequence 16, Appl1
18	1551	25.5	1102	15	US-10-185-867-36	Sequence 36, Appl1
19	1541.5	25.4	1083	15	US-10-160-224-1	Sequence 1, Appl1
20	1540.5	25.3	1083	10	US-09-965-830-2	Sequence 2, Appl1
21	1540.5	25.3	1083	15	US-10-185-867-2	Sequence 2, Appl1
22	1536	25.3	1082	15	US-10-121-746-20	Sequence 20, Appl1
23	1529.5	25.2	1080	9	US-09-119-855-2	Sequence 2, Appl1
24	1515.5	24.9	521	10	US-09-813-320-4	Sequence 4, Appl1
25	1511.5	24.9	530	10	US-09-813-320-2	Sequence 2, Appl1
26	1485.5	24.4	1284	9	US-09-119-855-11	Sequence 11, Appl1
27	1485.5	24.4	1284	15	US-10-160-224-9	Sequence 9, Appl1
28	1485.5	24.4	1284	15	US-10-185-867-11	Sequence 11, Appl1
29	1453.5	23.9	1174	15	US-10-174-613-7	Sequence 7, Appl1
30	1419	23.3	960	15	US-10-188-308-21	Sequence 21, Appl1
31	1419	23.3	960	15	US-10-188-296-21	Sequence 21, Appl1
32	1419	23.3	960	15	US-10-188-341-21	Sequence 21, Appl1
33	1419	23.3	960	15	US-10-188-297-21	Sequence 21, Appl1
34	1415	23.3	962	15	US-10-188-308-3	Sequence 3, Appl1
35	1415	23.3	962	15	US-10-188-296-3	Sequence 3, Appl1
36	1415	23.3	962	15	US-10-188-341-3	Sequence 3, Appl1
37	1415	23.3	962	15	US-10-188-297-3	Sequence 3, Appl1
38	1415	23.3	962	15	US-10-174-613-5	Sequence 5, Appl1
39	1409.5	23.2	962	15	US-10-174-613-4	Sequence 4, Appl1
40	1407	23.1	962	15	US-10-188-308-24	Sequence 24, Appl1
41	1407	23.1	962	15	US-10-188-296-24	Sequence 24, Appl1
42	1407	23.1	962	15	US-10-188-341-24	Sequence 24, Appl1
43	1407	23.1	962	15	US-10-188-297-24	Sequence 24, Appl1
44	1406.5	23.1	967	15	US-10-188-308-22	Sequence 22, Appl1
45	1406.5	23.1	987	15	US-10-188-296-22	Sequence 22, Appl1

ALIGNMENTS

RESULT 1
US-09-735-995-2
; Sequence 2, Application US/09735995
; Patent No. US20010034024A1
; GENERAL INFORMATION:
; APPLICANT: Keating, Mark T.
; APPLICANT: Splawski, Igor
; TITLE OF INVENTION: MUTATIONS IN AND GENOMIC STRUCTURE OF HERG - A LONG QT
; TITLE OF INVENTION: SYNDROME GENE
; FILE REFERENCE: 2323-136
; CURRENT APPLICATION NUMBER: US/09/735, 995
; CURRENT FILING DATE: 2000-12-14
; PRIOR APPLICATION NUMBER: 09/226, 012
; PRIOR FILING DATE: 1999-01-06
; NUMBER OF SEQ ID NOS: 116
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 2
; LENGTH: 1159
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-735-995-2

Query Match 100.0%; Score 6079; DB 9; Length 1159;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1159; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 MPVRGHVAPONTFLDTIIRKFEKOSRKFLIARVENCAYICNDGFCGLCYRAEVM 60
DB 1 MPVRGHVAPONTFLDTIIRKFEKOSRKFLIARVENCAYICNDGFCGLCYRAEVM 60
OY 61 QRECTDFLHGPTQRRRAAQIALGAEERVEIAFYRKDSCEFLCIVDVPVKNEDG 120
DB 61 QRECTDFLHGPTQRRRAAQIALGAEERVEIAFYRKDSCEFLCIVDVPVKNEDG 120
OY 121 AVIMFLINFEVMEKDMVGSPADDTNHRGPTSWLAPGRAKTRRLPALATARESSV 180
DB 121 AVIMFLINFEVMEKDMVGSPADDTNHRGPTSWLAPGRAKTRRLPALATARESSV 180

181 RSGGAGAGAGAVVVDLTPAAPSSSESLADEVTAMDNVAGLGAPEERALVGPSP 240
181 RSGGAGAGAGAVVVDLTPAAPSSSESLADEVTAMDNVAGLGAPEERALVGPSP 240
241 PRSAPGOLPSRAHSLNDASGSSCSLARTSRSCASVBRASSADDEARAGVLP 300
241 PRSAPGOLPSRAHSLNDASGSSCSLARTSRSCASVBRASSADDEARAGVLP 300
301 RHAATGAMHPLRSLNSTSDSLVRYRTISKIPQITLNFVDLKGDPFLASPTSDREIA 360
301 RHAATGAMHPLRSLNSTSDSLVRYRTISKIPQITLNFVDLKGDPFLASPTSDREIA 360
361 PKIERHTNVEKTYOVLISGADVLPEKKLOAPRIHRTILHYSPEKAWMDLILLVIY 420
361 PKIERHTNVEKTYOVLISGADVLPEKKLOAPRIHRTILHYSPEKAWMDLILLVIY 420
421 TAVFTPYSAALFKETEGEPATECGYACQPLAVVDLIVDIMEFVDLINFRTTYVNAE 480
421 TAVFTPYSAALFKETEGEPATECGYACQPLAVVDLIVDIMEFVDLINFRTTYVNAE 480
481 EVVSHPGRIAVHYFKGWFLLDMVAIIPEDLLIFSGSSEELIGLKTARLLRVYARKLD 540
481 EVVSHPGRIAVHYFKGWFLLDMVAIIPEDLLIFSGSSEELIGLKTARLLRVYARKLD 540
541 RYSEYGAAYLELLMCTPALIHMLACTIYAIAGNNEQPMDSRIGLHMLGQIGKPYNS 600
541 RYSEYGAAYLELLMCTPALIHMLACTIYAIAGNNEQPMDSRIGLHMLGQIGKPYNS 600
601 GLGSPSIKDKYVTLALYFTFSSLTSGVGNVSPNTNSEKIFISICVWLIGSLMYASIFGNVS 660
601 GLGSPSIKDKYVTLALYFTFSSLTSGVGNVSPNTNSEKIFISICVWLIGSLMYASIFGNVS 660
661 AIIORLYSGARHYTQMLRVRETFRHOIPNPLRQRLRETFRHOIHAMSTYNGIDMAVYKGF 720
661 AIIORLYSGARHYTQMLRVRETFRHOIPNPLRQRLRETFRHOIHAMSTYNGIDMAVYKGF 720
721 PECLQADICHLNLSLLOHCKPFRGATKGCCLRALAMKKTTHAPGDTLVHAGDILYALY 780
721 PECLQADICHLNLSLLOHCKPFRGATKGCCLRALAMKKTTHAPGDTLVHAGDILYALY 780
781 FISGSEIELRGDVAVALIGKNDIFGEPLNLVARGKSGNDVRAITYCDLKHIRDDLE 840
781 FISGSEIELRGDVAVALIGKNDIFGEPLNLVARGKSGNDVRAITYCDLKHIRDDLE 840
841 VLDMYPEFSDFHFWSSLEITFNLRDNTMIPGSPGSTELEGFSROKRLSFRRTDKDTE 900
841 VLDMYPEFSDFHFWSSLEITFNLRDNTMIPGSPGSTELEGFSROKRLSFRRTDKDTE 900
901 QPGEVSALGPRAGAGPSSRGRPGGPGWGESPPSSGSPESSEDEBPGHSSSPLRLVPSS 960
901 QPGEVSALGPRAGAGPSSRGRPGGPGWGESPPSSGSPESSEDEBPGHSSSPLRLVPSS 960
961 PRPGEPPGEPPLMEDCKSSDTCNPLSGAFSGVSNIFSGDSCRGRYOELPRCPAPTP 1020
961 PRPGEPPGEPPLMEDCKSSDTCNPLSGAFSGVSNIFSGDSCRGRYOELPRCPAPTP 1020
1021 SLNLIPLSSPPGRPRRGDYESRLDALQROLNRLLETRLSDMATVLOLQROMTLVPAYSA 1080
1021 SLNLIPLSSPPGRPRRGDYESRLDALQROLNRLLETRLSDMATVLOLQROMTLVPAYSA 1080
1081 VTTTGPGRGTSTPLPSPPLPTLTLDSISQVSOEMACBELPPGAPBELPQEGPTRRLSLPG 1140
1081 VTTTGPGRGTSTPLPSPPLPTLTLDSISQVSOEMACBELPPGAPBELPQEGPTRRLSLPG 1140
1141 QLGALTSOPLRHGSDPCS 1159
1141 QLGALTSOPLRHGSDPCS 1159

RESULT 2
US-09-735-995-4
; Sequence 4, Application US/09735995
; Patent No. US20010034024A1

GENERAL INFORMATION:
; APPLICANT: Keating, Mark T.
; APPLICANT: Splawski, Igor
; TITLE OF INVENTION: MUTATIONS IN AND GENOMIC STRUCTURE OF HERG - A LONG QT
; FILE REFERENCE: 2323-136
; CURRENT APPLICATION NUMBER: US/09/735, 995
; PRIOR FILING DATE: 2000-12-14
; PRIOR FILING DATE: 1999-01-06
; NUMBER OF SEQ ID NOS: 116
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 1159
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-735-995-4
Query Match 100.0%; Score 6079; DB 9; Length 1159;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1159; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
1 MPVRGHVAPONTFLDTIRKEGOSRKFITANARVENCAYICNDGFCELGYSRAEVM 60
1 MPVRGHVAPONTFLDTIRKEGOSRKFITANARVENCAYICNDGFCELGYSRAEVM 60
61 QRPCTCFLHGPRTORRAAOIAQALLGAEBRKYEAIFYRKDSCFLCLVDVPEVKEDG 120
61 QRPCTCFLHGPRTORRAAOIAQALLGAEBRKYEAIFYRKDSCFLCLVDVPEVKEDG 120
121 AVIMEFLNEFYVMKDMGSPAHDTNHRGPTSMAPRATFRKLKLPALLALTRARESSV 180
121 AVIMEFLNEFYVMKDMGSPAHDTNHRGPTSMAPRATFRKLKLPALLALTRARESSV 180
181 RSGGAGAGAGAVVVDLTPAAPSSSESLADEVTAMDNVAGLGAPEERALVGPSP 240
181 RSGGAGAGAGAVVVDLTPAAPSSSESLADEVTAMDNVAGLGAPEERALVGPSP 240
241 PRSAPGOLPSRAHSLNDASGSSCSLARTSRSCASVBRASSADDEARAGVLP 300
241 PRSAPGOLPSRAHSLNDASGSSCSLARTSRSCASVBRASSADDEARAGVLP 300
301 RHAATGAMHPLRSLNSTSDSLVRYRTISKIPQITLNFVDLKGDPFLASPTSDREIA 360
301 RHAATGAMHPLRSLNSTSDSLVRYRTISKIPQITLNFVDLKGDPFLASPTSDREIA 360
361 PKIERHTNVEKTYOVLISGADVLPEKKLOAPRIHRTILHYSPEKAWMDLILLVIY 420
361 PKIERHTNVEKTYOVLISGADVLPEKKLOAPRIHRTILHYSPEKAWMDLILLVIY 420
421 TAVFTPYSAALFKETEGEPATECGYACQPLAVVDLIVDIMEFVDLINFRTTYVNAE 480
421 TAVFTPYSAALFKETEGEPATECGYACQPLAVVDLIVDIMEFVDLINFRTTYVNAE 480
481 EVVSHPGRIAVHYFKGWFLLDMVAIIPEDLLIFSGSSEELIGLKTARLLRVYARKLD 540
481 EVVSHPGRIAVHYFKGWFLLDMVAIIPEDLLIFSGSSEELIGLKTARLLRVYARKLD 540
541 RYSEYGAAYLELLMCTPALIHMLACTIYAIAGNNEQPMDSRIGLHMLGQIGKPYNS 600
541 RYSEYGAAYLELLMCTPALIHMLACTIYAIAGNNEQPMDSRIGLHMLGQIGKPYNS 600
601 GLGSPSIKDKYVTLALYFTFSSLTSGVGNVSPNTNSEKIFISICVWLIGSLMYASIFGNVS 660
601 GLGSPSIKDKYVTLALYFTFSSLTSGVGNVSPNTNSEKIFISICVWLIGSLMYASIFGNVS 660
661 AIIORLYSGARHYTQMLRVRETFRHOIPNPLRQRLRETFRHOIHAMSTYNGIDMAVYKGF 720
661 AIIORLYSGARHYTQMLRVRETFRHOIPNPLRQRLRETFRHOIHAMSTYNGIDMAVYKGF 720
721 PECLQADICHLNLSLLOHCKPFRGATKGCCLRALAMKKTTHAPGDTLVHAGDILYALY 780
721 PECLQADICHLNLSLLOHCKPFRGATKGCCLRALAMKKTTHAPGDTLVHAGDILYALY 780

QY 781 FISRGSIEILRGDVVAALIGKNDIFGEPLNLVAPGKSNQDVRALTYCDLHKIHRDLE 840
DB 781 FISRGSIEILRGDVVAALIGKNDIFGEPLNLVAPGKSNQDVRALTYCDLHKIHRDLE 840
QY 841 VLDMPYEFSDHFWSSLEITFNLRDNTNIPGSPGSTELEGFSRQRRKLSFRRTDOTE 900
DB 841 VLDMPYEFSDHFWSSLEITFNLRDNTNIPGSPGSTELEGFSRQRRKLSFRRTDOTE 900
QY 901 QPGEVSLGPRGAGAGSSRGPGGPGWGESPPSSGESSEDEGPRSSSPLRLVPFSS 960
DB 901 QPGEVSLGPRGAGAGSSRGPGGPGWGESPPSSGESSEDEGPRSSSPLRLVPFSS 960
QY 961 PRPGEPGGEPLMEDCEKSDTCNPLSGAFSGVSNFTSFGDSRGQOYELPRCPATP 1020
DB 961 PRPGEPGGEPLMEDCEKSDTCNPLSGAFSGVSNFTSFGDSRGQOYELPRCPATP 1020
QY 1021 SLNIPPLSSPGRPRGDSRDLALQRLNRLLETRLSADNATVQLLQRLMTLVPAYSA 1080
DB 1021 SLNIPPLSSPGRPRGDSRDLALQRLNRLLETRLSADNATVQLLQRLMTLVPAYSA 1080
QY 1081 VTPPGGPTSTPLPSPPLPTLTLDSLSQVSOFMACEELPRGAPELPQGGPTRRLSLPG 1140
DB 1081 VTPPGGPTSTPLPSPPLPTLTLDSLSQVSOFMACEELPRGAPELPQGGPTRRLSLPG 1140
QY 1141 QLGALTSQPLHRHSDPGS 1159
DB 1141 QLGALTSQPLHRHSDPGS 1159

RESULT 3
US-09-119-855-10
; Sequence 10, Application US/09119855
; Patent No. US20020099197A1
; GENERAL INFORMATION:
; APPLICANT: Curtis, Roy A. J.
; TITLE OF INVENTION: NOVEL POTASSIUM CHANNEL MOLECULES AND USES THEREFOR
; FILE REFERENCE: mmi-055
; CURRENT APPLICATION NUMBER: US/09/119, 855
; CURRENT FILING DATE: 1998-07-21
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 10
; LENGTH: 1159
; TYPE: PRN
; ORGANISM: Homo sapiens
US-09-119-855-10

Query Match 100.0%; Score 6079; DB 9; Length 1159;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1159; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MPRRGHVADPONTFLDTIIRKEGQSRKFTIANARVENCAYIYCNDGFCGLGYSRAEVM 60
DB 1 MPRRGHVADPONTFLDTIIRKEGQSRKFTIANARVENCAYIYCNDGFCGLGYSRAEVM 60
QY 61 QPRTCTDFLHGPRTORRAAQIAQALIGAERKEVEIAFYRRDGSCLCLVDVYVKNEDG 120
DB 61 QPRTCTDFLHGPRTORRAAQIAQALIGAERKEVEIAFYRRDGSCLCLVDVYVKNEDG 120
QY 121 AVIMEFILNFEVVEKDWVGSPPADTNHGRPTSMVLAAGRAKTEFLKLPALLALTARSSV 180
DB 121 AVIMEFILNFEVVEKDWVGSPPADTNHGRPTSMVLAAGRAKTEFLKLPALLALTARSSV 180
QY 181 RSGGAGAGAPGAVVDVDTLTPAAPSESLALDEVTAMDNHVAAGLGAERERLALVFGSP 240
DB 181 RSGGAGAGAPGAVVDVDTLTPAAPSESLALDEVTAMDNHVAAGLGAERERLALVFGSP 240
QY 241 PRSAGGOLSPRAHSLNLPDASGSSCLAFRRSRESCASVRRASSADDIEMRAGVLPFP 300
DB 241 PRSAGGOLSPRAHSLNLPDASGSSCLAFRRSRESCASVRRASSADDIEMRAGVLPFP 300
QY 301 RHASTGAMHDLRGLNSTSDSLVRYRTISKIPQITLNFVDLKGDFLASPTSDREIIA 360

DB 301 RHASTGAMHDLRGLNSTSDSLVRYRTISKIPQITLNFVDLKGDFLASPTSDREIIA 360
QY 361 PKIKERHNTKEKVTQVLSIGADVLPYKIQAPRIHRWTLHYSPFAVADWMLILLVIY 420
DB 361 PKIKERHNTKEKVTQVLSIGADVLPYKIQAPRIHRWTLHYSPFAVADWMLILLVIY 420
QY 421 TAVFTPSAFLKETEPEGPRATECGYACOPLAVVDLIVDIMEFVILINFRTTYVANE 480
DB 421 TAVFTPSAFLKETEPEGPRATECGYACOPLAVVDLIVDIMEFVILINFRTTYVANE 480
QY 481 EVVSHPGRIAVHYFKGFLIDMVAALPFDDLIFGSGSEELIGLKTARLLRLVYARKLD 540
DB 481 EVVSHPGRIAVHYFKGFLIDMVAALPFDDLIFGSGSEELIGLKTARLLRLVYARKLD 540
QY 541 RYSEYGAVALFELMCFTALAHMLACIWAIGNNEQPHMDSRGWLNHLDQIGKPNSS 600
DB 541 RYSEYGAVALFELMCFTALAHMLACIWAIGNNEQPHMDSRGWLNHLDQIGKPNSS 600
QY 601 GLGSPSFKDKYVYALYFTFSSLSVSGGNVSPNTNSEKIFISICVMLIGSLMTASIFGNV 660
DB 601 GLGSPSFKDKYVYALYFTFSSLSVSGGNVSPNTNSEKIFISICVMLIGSLMTASIFGNV 660
QY 661 AIIQRLYSGTARYHTQMLRVREFIRHQIDNPRLQRLEEXYFOHAMSYTNGIDMANAVLKG 720
DB 661 AIIQRLYSGTARYHTQMLRVREFIRHQIDNPRLQRLEEXYFOHAMSYTNGIDMANAVLKG 720
QY 721 PECLQADICLHLNLSLQHRKPPRGATGCLRALAMKFTTHAPPGDTLVHAGDLTALY 780
DB 721 PECLQADICLHLNLSLQHRKPPRGATGCLRALAMKFTTHAPPGDTLVHAGDLTALY 780
QY 781 FISRGSIEILRGDVVAALIGKNDIFGEPLNLVAPGKSNQDVRALTYCDLHKIHRDLE 840
DB 781 FISRGSIEILRGDVVAALIGKNDIFGEPLNLVAPGKSNQDVRALTYCDLHKIHRDLE 840
QY 841 VLDMPYEFSDHFWSSLEITFNLRDNTNIPGSPGSTELEGFSRQRRKLSFRRTDOTE 900
DB 841 VLDMPYEFSDHFWSSLEITFNLRDNTNIPGSPGSTELEGFSRQRRKLSFRRTDOTE 900
QY 901 QPGEVSLGPRGAGAGSSRGPGGPGWGESPPSSGESSEDEGPRSSSPLRLVPFSS 960
DB 901 QPGEVSLGPRGAGAGSSRGPGGPGWGESPPSSGESSEDEGPRSSSPLRLVPFSS 960
QY 961 PRPGEPGGEPLMEDCEKSDTCNPLSGAFSGVSNFTSFGDSRGQOYELPRCPATP 1020
DB 961 PRPGEPGGEPLMEDCEKSDTCNPLSGAFSGVSNFTSFGDSRGQOYELPRCPATP 1020
QY 1021 SLNIPPLSSPGRPRGDSRDLALQRLNRLLETRLSADNATVQLLQRLMTLVPAYSA 1080
DB 1021 SLNIPPLSSPGRPRGDSRDLALQRLNRLLETRLSADNATVQLLQRLMTLVPAYSA 1080
QY 1081 VTPPGGPTSTPLPSPPLPTLTLDSLSQVSOFMACEELPRGAPELPQGGPTRRLSLPG 1140
DB 1081 VTPPGGPTSTPLPSPPLPTLTLDSLSQVSOFMACEELPRGAPELPQGGPTRRLSLPG 1140
QY 1141 QLGALTSQPLHRHSDPGS 1159
DB 1141 QLGALTSQPLHRHSDPGS 1159

RESULT 4
US-10-000-151b-3
; Sequence 3, Application US/10000151B
; Publication No. US20030013136A1
; GENERAL INFORMATION:
; APPLICANT: Balser, Jeffrey R.
; TITLE OF INVENTION: HUMAN K₁ REGULATION OF HERG POTASSIUM CHANNEL BLOCK
; FILE REFERENCE: Vanderbilt Ref No. US20030013136A1 VU0120; Attorney Docket No. US2
; CURRENT FILING DATE: 2000-10-30
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 3
; LENGTH: 1159
; TYPE: PRT
; ORGANISM: Homo sapiens
us-10-000-151b-3

```
Query Match      100.0%; Score 6079; DB 15; Length 1159;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1159; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 MPVRGHAAPONTFTDTIIRKEEGSRKFTIANARVENCAYICNDGFCGLGYSRAEVM 60
DB      1 MPVRGHAAPONTFTDTIIRKEEGSRKFTIANARVENCAYICNDGFCGLGYSRAEVM 60
QY      61 QRPCTCDEFLHGPRTQRRAAQIAQALLGAEEKVEIAFYRKDSCFCFLVDVVPVKNDG 120
DB      61 QRPCTCDEFLHGPRTQRRAAQIAQALLGAEEKVEIAFYRKDSCFCFLVDVVPVKNDG 120
QY      121 AVIMFILNEFVVMKDMVSPAHDTNHRGPPPTSMILAPGRAKTFRLKLTALLATARESSV 180
DB      121 AVIMFILNEFVVMKDMVSPAHDTNHRGPPPTSMILAPGRAKTFRLKLTALLATARESSV 180
QY      181 RSGGAGAGAGAAVVDVLTTPAASSESLADEVTAMDNDNVAGLGPAAEERRALVGPSP 240
DB      181 RSGGAGAGAGAAVVDVLTTPAASSESLADEVTAMDNDNVAGLGPAAEERRALVGPSP 240
QY      241 PRSAPGOLPSRAHSLNDPASGSSCSLARTSRSCASVRRASADIEAMRAGVLPPPP 300
DB      241 PRSAPGOLPSRAHSLNDPASGSSCSLARTSRSCASVRRASADIEAMRAGVLPPPP 300
QY      301 RHASTGAMHPLRSGLNSTSDSLVRYRTISKIPQITLNFVDLKGDPFLASPTSDREITIA 360
DB      301 RHASTGAMHPLRSGLNSTSDSLVRYRTISKIPQITLNFVDLKGDPFLASPTSDREITIA 360
QY      361 PRIKERTHNVEKYOVLSIGADVLPEYKLOAPRIHRMTILHYSPEKAVMDLILLVIY 420
DB      361 PRIKERTHNVEKYOVLSIGADVLPEYKLOAPRIHRMTILHYSPEKAVMDLILLVIY 420
QY      421 TAVFTPYSAANFLKTEEGPPATCGYACQPLAVVDLIVDIMEFVLDLINFRTTYVANE 480
DB      421 TAVFTPYSAANFLKTEEGPPATCGYACQPLAVVDLIVDIMEFVLDLINFRTTYVANE 480
QY      481 EVSHPGRIAVHYFGWFLIDMVAAIPBDLLIFGSGSEELIGLKTARLLRYVARAKLD 540
DB      481 EVSHPGRIAVHYFGWFLIDMVAAIPBDLLIFGSGSEELIGLKTARLLRYVARAKLD 540
QY      541 RYSEYGAANFLKCTEFALIAHMLACIWAIGNMOPIMDSRIGLHMLGDOIGKPYNS 600
DB      541 RYSEYGAANFLKCTEFALIAHMLACIWAIGNMOPIMDSRIGLHMLGDOIGKPYNS 600
QY      601 GLGGSIKDKYVTALYFTFSSLTSGVGNVSPNTNSEKIFSIQVMLIGSLMYASIFGNVS 660
DB      601 GLGGSIKDKYVTALYFTFSSLTSGVGNVSPNTNSEKIFSIQVMLIGSLMYASIFGNVS 660
QY      661 AIIQRLVSGTARHTOMLRVREFIRFHOIPNPRLQRLEEFYFOHAMSTYNGIDMAVLCGF 720
DB      661 AIIQRLVSGTARHTOMLRVREFIRFHOIPNPRLQRLEEFYFOHAMSTYNGIDMAVLCGF 720
QY      721 PECLADICLHNSLLOHCKPFRGATGCLRALAMKKTTHAPGDTLVNAGDILLALY 780
DB      721 PECLADICLHNSLLOHCKPFRGATGCLRALAMKKTTHAPGDTLVNAGDILLALY 780
QY      781 FISRGSIEILRGDVVAALIGKNDIFGEPLNYARPGKSGNDVRLATYCDLKHIRHDDLE 840
DB      781 FISRGSIEILRGDVVAALIGKNDIFGEPLNYARPGKSGNDVRLATYCDLKHIRHDDLE 840
QY      841 VLDWYPRSDHFWSSLETETENLRTNMTFGSPSTLEBGGESRQKRLSTRRRDKYTE 900
DB      841 VLDWYPRSDHFWSSLETETENLRTNMTFGSPSTLEBGGESRQKRLSTRRRDKYTE 900
QY      901 QPGEVSALGPGRAGPSSRGRPGPWPGEPSGSPSSGSESEDEGPGSSPLRLVPPSS 960
DB      901 QPGEVSALGPGRAGPSSRGRPGPWPGEPSGSPSSGSESEDEGPGSSPLRLVPPSS 960
```

```
QY      961 PRPGEPPGGEPLMEDCKSSDTCNPLSGASGVSNIFSWGDSNGROYELPRCPAPT 1020
DB      961 PRPGEPPGGEPLMEDCKSSDTCNPLSGASGVSNIFSWGDSNGROYELPRCPAPT 1020
QY      1021 SLNITPLSSPGRRRPPGDVESRLDALQROLNRLERFLSDMATVLOLQROMTLYPPAYSA 1080
DB      1021 SLNITPLSSPGRRRPPGDVESRLDALQROLNRLERFLSDMATVLOLQROMTLYPPAYSA 1080
QY      1081 VTPGPGPTSTSPILPVPPLPTLTLDLSQVSOFAWCELPFGAELPQEGPTRRLSLPG 1140
DB      1081 VTPGPGPTSTSPILPVPPLPTLTLDLSQVSOFAWCELPFGAELPQEGPTRRLSLPG 1140
QY      1141 QLGALTSQPLHRHGSPPGS 1159
DB      1141 QLGALTSQPLHRHGSPPGS 1159

RESULT 5
US-10-193-692-5
; Sequence 5, Application US/10193692
; Publication No. US20030074682A1
; GENERAL INFORMATION:
; APPLICANT: Exelixis Deutschland GmbH
; TITLE OF INVENTION: ISOLATION, CHARACTERIZATION, AND USE OF A NOVEL TELEOST POTASS
; FILE REFERENCE: AR02-005C
; CURRENT APPLICATION NUMBER: US/10/193,692
; CURRENT FILING DATE: 2002-07-11
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
; LENGTH: 1159
; TYPE: PRT
; ORGANISM: Homo sapiens
us-10-193-692-5

Query Match      100.0%; Score 6079; DB 15; Length 1159;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1159; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 MPVRGHAAPONTFTDTIIRKEEGSRKFTIANARVENCAYICNDGFCGLGYSRAEVM 60
DB      1 MPVRGHAAPONTFTDTIIRKEEGSRKFTIANARVENCAYICNDGFCGLGYSRAEVM 60
QY      61 QRPCTCDEFLHGPRTQRRAAQIAQALLGAEEKVEIAFYRKDSCFCFLVDVVPVKNDG 120
DB      61 QRPCTCDEFLHGPRTQRRAAQIAQALLGAEEKVEIAFYRKDSCFCFLVDVVPVKNDG 120
QY      121 AVIMFILNEFVVMKDMVSPAHDTNHRGPPPTSMILAPGRAKTFRLKLTALLATARESSV 180
DB      121 AVIMFILNEFVVMKDMVSPAHDTNHRGPPPTSMILAPGRAKTFRLKLTALLATARESSV 180
QY      181 RSGGAGAGAGAAVVDVLTTPAASSESLADEVTAMDNDNVAGLGPAAEERRALVGPSP 240
DB      181 RSGGAGAGAGAAVVDVLTTPAASSESLADEVTAMDNDNVAGLGPAAEERRALVGPSP 240
QY      241 PRSAPGOLPSRAHSLNDPASGSSCSLARTSRSCASVRRASADIEAMRAGVLPPPP 300
DB      241 PRSAPGOLPSRAHSLNDPASGSSCSLARTSRSCASVRRASADIEAMRAGVLPPPP 300
QY      301 RHASTGAMHPLRSGLNSTSDSLVRYRTISKIPQITLNFVDLKGDPFLASPTSDREITIA 360
DB      301 RHASTGAMHPLRSGLNSTSDSLVRYRTISKIPQITLNFVDLKGDPFLASPTSDREITIA 360
QY      361 PRIKERTHNVEKYOVLSIGADVLPEYKLOAPRIHRMTILHYSPEKAVMDLILLVIY 420
DB      361 PRIKERTHNVEKYOVLSIGADVLPEYKLOAPRIHRMTILHYSPEKAVMDLILLVIY 420
QY      421 TAVFTPYSAANFLKTEEGPPATCGYACQPLAVVDLIVDIMEFVLDLINFRTTYVANE 480
DB      421 TAVFTPYSAANFLKTEEGPPATCGYACQPLAVVDLIVDIMEFVLDLINFRTTYVANE 480
QY      481 EVSHPGRIAVHYFGWFLIDMVAAIPBDLLIFGSGSEELIGLKTARLLRYVARAKLD 540
```

|||||
Db 481 EVVSHPRIRIAVHYFKGFFLIDMVAIIPFDLIFGSGSEELIGLTKTARLLRLVAVARKLD 540
QY 541 RYSEYGAAYFLMLCTFALLAHMLACTIWAIAIGNNEQHMDSRIGMLNLDQICKPNSS 600
Db 541 RYSEYGAAYFLMLCTFALLAHMLACTIWAIAIGNNEQHMDSRIGMLNLDQICKPNSS 600
QY 601 GLGSPSIRKDYVTALYTFSSLSVSGFNVSPTNNSKIFSIQVMLIGSLMYASIFGNVS 660
Db 601 GLGSPSIRKDYVTALYTFSSLSVSGFNVSPTNNSKIFSIQVMLIGSLMYASIFGNVS 660
QY 661 AIIORLYSGTARHTOMLRVREFIRFHOIPNPLRQRLLEYFQOHAMSTYNGIDMAVYKGF 720
Db 661 AIIORLYSGTARHTOMLRVREFIRFHOIPNPLRQRLLEYFQOHAMSTYNGIDMAVYKGF 720
QY 721 PECLQADICHLNLSLQHKPFPGATKGCIRALAMFKTTTHAPPGDTLVHAGDLTALY 780
Db 721 PECLQADICHLNLSLQHKPFPGATKGCIRALAMFKTTTHAPPGDTLVHAGDLTALY 780
QY 781 FISGSIIEILRGDVVAAILKNDIFGEPLMIYARPGKSGDVRLATYCDLHKIHRDLE 840
Db 781 FISGSIIEILRGDVVAAILKNDIFGEPLMIYARPGKSGDVRLATYCDLHKIHRDLE 840
QY 841 VLDMPFSDHFWSSLEITFNLKDTNMI PGSPGSTELEGGFSRQRRKLSFRRTDKOTE 900
Db 841 VLDMPFSDHFWSSLEITFNLKDTNMI PGSPGSTELEGGFSRQRRKLSFRRTDKOTE 900
QY 901 QPGEVSAALGPRAGAGSSRGRCGPWGESPSSPSESEDEGPRSSSPLRLVPFSS 960
Db 901 QPGEVSAALGPRAGAGSSRGRCGPWGESPSSPSESEDEGPRSSSPLRLVPFSS 960
QY 961 PRPGEPGGEPLMEDEKSSDTCNPLSGAFSGVSNIFSWGDSRGROYELPRCPAPTP 1020
Db 961 PRPGEPGGEPLMEDEKSSDTCNPLSGAFSGVSNIFSWGDSRGROYELPRCPAPTP 1020
QY 1021 SLNIPILSSGRRRGVDSRDLAOLQRLNLETRLSADMATYVQLQRLQMTLVPPAYSA 1080
Db 1021 SLNIPILSSGRRRGVDSRDLAOLQRLNLETRLSADMATYVQLQRLQMTLVPPAYSA 1080
QY 1081 VTPPGPSTSPILPVSPILTLTLDLSQVSOFMACEELPRGAPELPOGCPTRRLSLPG 1140
Db 1081 VTPPGPSTSPILPVSPILTLTLDLSQVSOFMACEELPRGAPELPOGCPTRRLSLPG 1140
QY 1141 QLGALTSQPLHRHGSDDPGS 1159
Db 1141 QLGALTSQPLHRHGSDDPGS 1159

RESULT 6
US-10-185-867-10
: Sequence 10, Application US/10185867
: Publication No. US20030104429A1
: GENERAL INFORMATION:
: APPLICANT: Curtis, Rory A.J.
: TITLE OF INVENTION: NOVEL POTASSIUM CHANNEL MOLECULES AND USES THEREFOR
: FILE REFERENCE: LMI-055CP
: CURRENT APPLICATION NUMBER: US/10/185, 867
: PRIOR FILING DATE: 1999-07-21
: PRIOR APPLICATION NUMBER: US/09/358, 383
: PRIOR FILING DATE: 1998-07-21
: NUMBER OF SEQ ID NOS: 36
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 10
: LENGTH: 1159
: TYPE: PRT
: ORGANISM: Homo sapiens
US-10-185-867-10

Query Match 100.0%; Score 6079; DB 15; Length 1159;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1159; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MPVRGHVAPONTFLDTIIRKEGOSRKFIITANARVENCAYIYCNDFCELCGYSRAEVM 60
Db 1 MPVRGHVAPONTFLDTIIRKEGOSRKFIITANARVENCAYIYCNDFCELCGYSRAEVM 60
QY 61 QRPCTDFLHGPRTORRAAAQIAQALLGAERKEVIAFYRRKDGSCFLVDVVPVKNEG 120
Db 61 QRPCTDFLHGPRTORRAAAQIAQALLGAERKEVIAFYRRKDGSCFLVDVVPVKNEG 120
QY 121 AVIMFIINFEVYMKDWGSPAHDTNHRGPTSLAAGRAKTFPLKLPALLALTARBSV 180
Db 121 AVIMFIINFEVYMKDWGSPAHDTNHRGPTSLAAGRAKTFPLKLPALLALTARBSV 180
QY 181 RSGGAGAGAPGAVVVDLTPAAPSSESLALDEVTAMDNDHVAGLGAEEERATLVGSP 240
Db 181 RSGGAGAGAPGAVVVDLTPAAPSSESLALDEVTAMDNDHVAGLGAEEERATLVGSP 240
QY 241 PRSAPGOLPSBRAHSLNPDASGSSCLARTSRSCASVRRASSADDIEAMRAGVLPBP 300
Db 241 PRSAPGOLPSBRAHSLNPDASGSSCLARTSRSCASVRRASSADDIEAMRAGVLPBP 300
QY 301 RHASTGAMHPLRSGLSNSTSDSLVYRTITSKIPQITLNFYDLKGFPLASPTSDREIA 360
Db 301 RHASTGAMHPLRSGLSNSTSDSLVYRTITSKIPQITLNFYDLKGFPLASPTSDREIA 360
QY 361 PKIKERTHNTKTYOVLISGADVLPYKLOAPRIHMTILHVSFKAAMDMLLLVIY 420
Db 361 PKIKERTHNTKTYOVLISGADVLPYKLOAPRIHMTILHVSFKAAMDMLLLVIY 420
QY 421 TAVETPYSAFLKETEGBEPATECGYACOPLAVIDLVIDIMFIVDLINERTTYVANE 480
Db 421 TAVETPYSAFLKETEGBEPATECGYACOPLAVIDLVIDIMFIVDLINERTTYVANE 480
QY 481 EYVSHPRIRIAVHYFKGFFLIDMVAIIPFDLIFGSGSEELIGLTKTARLLRLVAVARKLD 540
Db 481 EYVSHPRIRIAVHYFKGFFLIDMVAIIPFDLIFGSGSEELIGLTKTARLLRLVAVARKLD 540
QY 541 RYSEYGAAYFLMLCTFALLAHMLACTIWAIAIGNNEQHMDSRIGMLNLDQICKPNSS 600
Db 541 RYSEYGAAYFLMLCTFALLAHMLACTIWAIAIGNNEQHMDSRIGMLNLDQICKPNSS 600
QY 601 GLGSPSIRKDYVTALYTFSSLSVSGFNVSPTNNSKIFSIQVMLIGSLMYASIFGNVS 660
Db 601 GLGSPSIRKDYVTALYTFSSLSVSGFNVSPTNNSKIFSIQVMLIGSLMYASIFGNVS 660
QY 661 AIIORLYSGTARHTOMLRVREFIRFHOIPNPLRQRLLEYFQOHAMSTYNGIDMAVYKGF 720
Db 661 AIIORLYSGTARHTOMLRVREFIRFHOIPNPLRQRLLEYFQOHAMSTYNGIDMAVYKGF 720
QY 721 PECLQADICHLNLSLQHKPFPGATKGCIRALAMFKTTTHAPPGDTLVHAGDLTALY 780
Db 721 PECLQADICHLNLSLQHKPFPGATKGCIRALAMFKTTTHAPPGDTLVHAGDLTALY 780
QY 781 FISGSIIEILRGDVVAAILKNDIFGEPLMIYARPGKSGDVRLATYCDLHKIHRDLE 840
Db 781 FISGSIIEILRGDVVAAILKNDIFGEPLMIYARPGKSGDVRLATYCDLHKIHRDLE 840
QY 841 VLDMPFSDHFWSSLEITFNLKDTNMI PGSPGSTELEGGFSRQRRKLSFRRTDKOTE 900
Db 841 VLDMPFSDHFWSSLEITFNLKDTNMI PGSPGSTELEGGFSRQRRKLSFRRTDKOTE 900
QY 901 QPGEVSAALGPRAGAGSSRGRCGPWGESPSSPSESEDEGPRSSSPLRLVPFSS 960
Db 901 QPGEVSAALGPRAGAGSSRGRCGPWGESPSSPSESEDEGPRSSSPLRLVPFSS 960
QY 961 PRPGEPGGEPLMEDEKSSDTCNPLSGAFSGVSNIFSWGDSRGROYELPRCPAPTP 1020
Db 961 PRPGEPGGEPLMEDEKSSDTCNPLSGAFSGVSNIFSWGDSRGROYELPRCPAPTP 1020
QY 1021 SLNIPILSSGRRRGVDSRDLAOLQRLNLETRLSADMATYVQLQRLQMTLVPPAYSA 1080
Db 1021 SLNIPILSSGRRRGVDSRDLAOLQRLNLETRLSADMATYVQLQRLQMTLVPPAYSA 1080

Qy	581	RIGIMLHNLGQIOKPKPY--SSGIGGSPSKDKYVALATFTFSSLSVSGGNSPMTNSKI	633
Db	583	DKIGLISLGLGIGIRNDSOSSGSPKDKYVALATFTFSSLSVSGGNSPMTNSKI	642
Qy	640	FSICVMLGSLMVASIFGNSAIIQIORYSGARVHTOMLRREFIRFQIPIPLRLEE	699
Db	643	FSICVMLGSLMVASIFGNSAIIQIORYSGARVHTOMLRKEFIRFQIPIPLRLEE	702
Qy	700	YFOHAMSTNGCIDMANVALKGPPECLQADICLHNLRSLOHCKPRPGATKGLRLAMKFK	759
Db	703	YFOHAMTYTNGIDMANVMVLKGPPECLQADICLHNLQONKARFGASKGLRLAMKFK	762
Qy	760	TTNHPGGTIVHAGDILVALFYIRNGSTIELRGVYVVAIIILKNDFGLRLVYRPGKS	819
Db	763	TTNHPGGTIVHAGDILVALFYIRNGSTIELKDIYVAIIILKNDFGEMVHLLYAKPGKS	822
Qy	820	GDVALATYCDLHKITHRDDLEVLDMYRPEFSDHFWSSLEITPNLRDNTMIPGSPSTLEEG	879
Db	823	ADVALATYCDLHKITHRDDLEVLDMYRPEFSDHFTLEITPNLRHES--AKSOSINDEG	880
Qy	880	GFSQORRKKLSFRRRTDKDTQRPGEVALGPRGAGSS--RGRGPGWGESPPSGR	935
Db	881	DTCKLRRRRLSFESEBDDFSKENSAN-----DADSDTIRRYOSKKHFEKKSRS	933
Qy	936	SSPESSEDEGP-----GRSSPLRLVPPSSPRPGE--PPGG-----EPLMEDEKSSDT	983
Db	934	SSPSSSIDDECKPFLGLGYVDSTPRMVYKAS--RHNGEELAPPSGNIHNDKRSCHKDITDT	991
Qy	984	-----CNPLSGAFSGVSNISFSGWSGRGOYELPRCPAPYPSLLNIPLSS	1029
Db	992	HSWERENARAQPERECSP-----SGIQR--AAWGISE-----TESDILTY-----	1027
Qy	1030	PGRRRGVGESERLDALORLNRLEPRLSADATVLIOLROMTVPRYSAVT-----	1082
Db	1028	-----GEVEGRDLIDQLRLKRESOMTTDIOAILDLOKOTVVPYPSVNTAGAEYOR	1081
Qy	1083	-----TPGP-----GPTSTSP-----LLPVSPLTLTLDLSIQ	1110
Db	1082	PILRLRTSHPRASIKTDRSPSPSQCEPFLDELSKLKSKESJSSGKRLMTASBDNLT	1141
Qy	1111	V--SOFMACELPRGA-----PRLPDEGPRRLSLPGQLGALTSQPLHRRGSDPG	1158
Db	1142	LKQDSDASSELDPQRKSYLHPRIHPSLPDSS--LSTVGILG-----LHRHVSDPG	1191
RESUT 10			
US-10-255-532-2			
; Sequence 2, Application US/10255532			
; Publication No. US2003009991A1			
; GENERAL INFORMATION:			
; APPLICANT: Silos-Santiago, Immaculada			
; TITLE OF INVENTION: METHODS OF USING 33751, A HUMAN			
; FILE REFERENCE: MP101-242P1RM			
; CURRENT APPLICATION NUMBER: US/10/255, 532			
; CURRENT FILING DATE: 2002-09-26			
; PRIOR APPLICATION NUMBER: 60/325,854			
; PRIOR FILING DATE: 2001-09-27			
; NUMBER OF SEQ ID NOS: 7			
; SOFTWARE: FastSeq for Windows Version 4.0			
; SEQ ID NO 2			
; LENGTH: 1196			
; TYPE: PRT			
; ORGANISM: Homo sapien			
US-10-255-532-2			
Query Match			
Best Local Similarity 54.8%, Score 3255; DB 15; Length 1196;			
Matches 700; Conservative 126; Mismatches 246; Indels 206; Gaps 25;			
Qy	1	MPVRGRVAPONTFLDTIRKFEQOSKRFTIYANAVENCAVITYCNDGFCGLGYSRAEVM	60

D6	1	HPVRGHGAPONTELGITIIIRREFEGCNKFFIIANAIVQACIIIIYCDNGFCENTGFSRPDVM	60
QY	61	QPECTCDFLHGPRIOORRAAOIAOALGAEERKEIAFYRKDGSCFLVDVVPYKKNEDG	120
D6	61	OKPCTCDLHGPETGRHIDIAOALLGSEERKEVYIYHKNGSPFICNTNIIIPKKNEDG	120
QY	121	AVIMILNEFVYEMEDMGSPAHDNNHGCPPTSWLAPRATFRKLALLATLTKRESSY	180
D6	121	VAMMIIINEFYEDNENATP-----EYVNPLPIKTYVNRKFRFGKFFGLRLVYLRKOSL	175
QY	181	RSKGAGAGAAVYVDVLTTPAAPSSESLADEVTAMDNVNAGGPPAEERRALVPG--	238
D6	176	POED-----PDVVID-----SKRSDSVAMKHKSPFKSCSPSEADPTKALQPSKC	225
QY	239	SPRSPAGOL-----PSPRAHSLNDPAGSSCSCLARTRSRESCAVYRARSADIDFAMRAG	294
D6	226	SPLVNSGPLDHSSEKROWMLPYDMLQSSOLSHSRRESLCSIRRASVHIDEGF--G	283
QY	235	VLP-----PPRNAS-----TGAMHPLRSGSLNSTRSDSLRYRPRISAIPOITLNEFVCLK	344
D6	284	VHPKNIFDRHASEDNGNRNVKGPFEHNHKLSSLGSTSDSLNKYSTINKIPOTLTFSEVK	343
QY	345	GDPFLAS-PTSDREIIARIKERTNMYEKTQVYLSLADAVLPETKLOAPRIHWTILHY	403
D6	344	TEKNSSPSSDCKTILAKVDRHTNVTAEKTYOVLSTGADVLPXKLTQPIRINKFTLIHY	403
QY	404	SPFKAVMDMLLLLVYTAFTVPYSAAFLKTEBGPATCGYACQPLAVVDLVIDMF	463
D6	404	SPFKAVMDMLLLLVYTAFTVPYSAAFLNDREE--OKRREGYSCSPLNVVDLVIDMF	462
QY	464	IVDLINERTYVYVANEVSHPGIAVHYKRGWFLDMVAAPDILLIFSGSSEE--L	520
D6	463	IIDLINERTYVYVONEEVSDPAITAIHYRGMFLDMVAAPDILLIFSGSDETTTL	522
QY	521	IGLKTARLRLVYRARKLIDRYSEKGAVFLMLCTFALIAHMLACIYWAIGNMDOPHMD	580
D6	523	IGLKTARLRLVYRARKLIDRYSEKGAVFLMLSMCIPALNHMLACIYWAIGNVERPYLT	582
QY	581	SRIGMLNHLGDOIGKPYN--SSGLGSPSTIKDKYVATLYTFSSLSVFGGNVSPNTNSKI	639
D6	583	DKIGLMDLSGOOIGKRYNDSOSSSGSPSTIKDKYVATLYTFSSLSVFGGNVSPNTNSKI	642
QY	640	FSICMGLISLMYASIFGNVSALIORLYSGARYTOMLRYREFIRFHOIPNPLORLEE	699
D6	643	FSICMGLISLMYASIFGNVSALIORLYSGARYHOMLRKEFIRFHOIPNPLORLEE	702
QY	700	YFOHASTYNGIDMNAVLKPEPCLOADICHLNLSLQHKPFGRAGKGCRLALAMFK	759
D6	703	YFOHASTYNGIDMNAVLKPEPCLOADICHLNLTLLQNKAFGAGKGCRLALAMFK	762
QY	760	TTAHAPGDTLVHAGDILLTALYFISRGSIEILRGDVVAIILKNDIIGEPPLNLYARPKSN	819
D6	763	TTAHAGDTLVHAGDILLTALYFISRGSIEISKNDVVAIILKNDIFGMVHLYARPKSN	822
QY	820	GDVALTYCDLHKIHRDILLVDMYPEPSHMSLITFNLRPTNM---IPGSPGSTE	876
D6	823	ADVALTYCDLHKIQREDLLEVDMPPEPSHFTLNLETFELTRHESAKADLLRQSMND	882
QY	877	LEGFSROKRRKILSFRRRTDQTEPGEVASALGPAGAGSSNRPGGPMGSESSGPS	936
D6	883	SEGDMCKLRKRLSESGEKE-----NSTN	908
QY	937	SPESSED-----EGPGRSSPLRLVPESSPRPGEPGPELMEDEKSSDT	983
D6	909	DPEDSADTIRHYQSSKRHFEEKRSSEFISII-----DDE	944
QY	964	CNPLSGASGVSNIFSPMGDSRGRYOELPCRPATPBLNIPL-----	1022
D6	945	OKPL---FSGIYDSSPGIGAKASGLDTEE---TYPTSGRMHIDKRSKCDITDMRSMER	997
QY	1028	SSPGRRRP-----GDVSRDLALOROLNRLLETRLSAMATVYL	1064
D6	998	ENAHQPEDSSSALQORAMGISTESDLYTGEVQRIDLLQOLNRLRESQWTTDIOQTL	1055

Matches 670; Conservative 85; Mismatches 153; Indels 329; Gaps 32;

QY 1 MPVRGHVAPONTFLDTIIRKFEKGSRKFIANARVENCAYICNDGCELCGYSRAEYM 60
 DB 1 MPVRGHVAPONTFLDTIIRKFEKGSRKFIANARVENCAYICNDGCELCGYSRAEYM 60
 QY 61 GPRCTDFLHGRTQRRAAQIAOALLGAEEKVEIATFRKQSGCFCLVDVVPKKNEDG 120
 DB 61 GPRCTDFLHGRTQRRAAQIAOALLGAEEKVEIATFRKQSGCFCLVDVVPKKNEDG 120
 QY 121 AVIMEFLNEEVMKDMVGSPPADTNHRGPTSWLAPGRAKTFRLKLPALLA-LTARESS 179
 DB 121 AVIMEFLNEEVMKDMVGSPPADTNHRGPTSWLAPGRAKTFRLKLPALLA-LTARESS 179
 QY 180 VR-----SGAGAGAGAGAVVVDVLPAPAPSSSLALDEVYAMDNDHVAAGLGAPEERRA 233
 DB 180 VR-----SGAGAGAGAGAVVVDVLPAPAPSSSLALDEVYAMDNDHVAAGLGAPEERRA 233
 QY 146 QRLLSQSFSGSESGHGRPG----- 164
 DB 146 QRLLSQSFSGSESGHGRPG----- 164
 QY 234 LVGPSPRRAPGOLPSPRAHSLNPDASGSSCLARTSRRESCASVRRASADIEAMRA 293
 DB 234 LVGPSPRRAPGOLPSPRAHSLNPDASGSSCLARTSRRESCASVRRASADIEAMRA 293
 QY 165 --GPG-----PG----- 169
 DB 165 --GPG-----PG----- 169
 QY 294 GVLPPPPRHASTGAMHPLRSGLLNSTSDLVRYRTISKIPQITLNFVLDKGPFLASPT 353
 DB 294 GVLPPPPRHASTGAMHPLRSGLLNSTSDLVRYRTISKIPQITLNFVLDKGPFLASPT 353
 QY 170 -----TGR-----GKRTISOIQTPLNFVEFNLEKRRSST 201
 DB 170 -----TGR-----GKRTISOIQTPLNFVEFNLEKRRSST 201
 QY 354 SDRETIAP-KIKERTHNTEKTVQVLSIGADVLPYKLOAPRIHRWTILHYSPPKAVMDW 412
 DB 354 SDRETIAP-KIKERTHNTEKTVQVLSIGADVLPYKLOAPRIHRWTILHYSPPKAVMDW 412
 QY 202 TEIEIIPAKHYVERQNTVEKTVQVLSIGADVLPYKLOAPRIHRWTILHYSPPKAVMDW 261
 DB 202 TEIEIIPAKHYVERQNTVEKTVQVLSIGADVLPYKLOAPRIHRWTILHYSPPKAVMDW 261
 QY 413 LILLVITYAVFTPSAFLKETEGRPATCEGACQAPLAVVDLIVDIMEFVDILINR 472
 DB 413 LILLVITYAVFTPSAFLKETEGRPATCEGACQAPLAVVDLIVDIMEFVDILINR 472
 QY 262 LILLVITYAVFTPSAFLKETEGRPATCEGACQAPLAVVDLIVDIMEFVDILINR 320
 DB 262 LILLVITYAVFTPSAFLKETEGRPATCEGACQAPLAVVDLIVDIMEFVDILINR 320
 QY 473 TTYVANEEVSHPRGRIAVHYFKGFWFLIDMAIPEFDLIFGSGSEE--LIGLKTARL 529
 DB 473 TTYVANEEVSHPRGRIAVHYFKGFWFLIDMAIPEFDLIFGSGSEE--LIGLKTARL 529
 QY 321 TTYVMTNEGVSHPRRIAVHYFKGFWFLIDMAIPEFDLIFGSGSEE--LIGLKTARL 380
 DB 321 TTYVMTNEGVSHPRRIAVHYFKGFWFLIDMAIPEFDLIFGSGSEE--LIGLKTARL 380
 QY 530 LRLVAVARKLDRISEYGAALFLMCTFALIAHMLACIYAGNVERPYLEHKGIMWLSL 589
 DB 530 LRLVAVARKLDRISEYGAALFLMCTFALIAHMLACIYAGNVERPYLEHKGIMWLSL 589
 QY 381 LRLVAVARKLDRISEYGAALFLMCTFALIAHMLACIYAGNVERPYLEHKGIMWLSL 440
 DB 381 LRLVAVARKLDRISEYGAALFLMCTFALIAHMLACIYAGNVERPYLEHKGIMWLSL 440
 QY 590 GJOICKPPNVSNG-LGGPSIKDKYVVAIVFTFSSLSVCGNVSPTNNSKISICVMILIG 648
 DB 590 GJOICKPPNVSNG-LGGPSIKDKYVVAIVFTFSSLSVCGNVSPTNNSKISICVMILIG 648
 QY 441 GJOICKPPNVSNG-LGGPSIKDKYVVAIVFTFSSLSVCGNVSPTNNSKISICVMILIG 500
 DB 441 GJOICKPPNVSNG-LGGPSIKDKYVVAIVFTFSSLSVCGNVSPTNNSKISICVMILIG 500
 QY 649 SLMTASIFGNVSAIIQRLYSGTARYHTQMLRVREFIRFHOIIPNPLRORLEEFQOHAWSY 708
 DB 649 SLMTASIFGNVSAIIQRLYSGTARYHTQMLRVREFIRFHOIIPNPLRORLEEFQOHAWSY 708
 QY 501 SLMTASIFGNVSAIIQRLYSGTARYHTQMLRVREFIRFHOIIPNPLRORLEEFQOHAWSY 560
 DB 501 SLMTASIFGNVSAIIQRLYSGTARYHTQMLRVREFIRFHOIIPNPLRORLEEFQOHAWSY 560
 QY 709 NGIDMNAVLKGFPECLQADICLHLNRSLLQHCPRFGATKGLRALAMKFTTHAPRGDT 768
 DB 709 NGIDMNAVLKGFPECLQADICLHLNRSLLQHCPRFGATKGLRALAMKFTTHAPRGDT 768
 QY 561 NGIDMNAVLKGFPECLQADICLHLNRSLLQHCPRFGATKGLRALAMKFTTHAPRGDT 620
 DB 561 NGIDMNAVLKGFPECLQADICLHLNRSLLQHCPRFGATKGLRALAMKFTTHAPRGDT 620
 QY 769 LVHADDLTALYFISRGSEIIRGQVVAAILCKNDIFGEPFLYLARPGKSNQDVALTYC 828
 DB 769 LVHADDLTALYFISRGSEIIRGQVVAAILCKNDIFGEPFLYLARPGKSNQDVALTYC 828
 QY 621 LVHADDLTALYFISRGSEIIRGQVVAAILCKNDIFGEPFLYLARPGKSNQDVALTYC 680
 DB 621 LVHADDLTALYFISRGSEIIRGQVVAAILCKNDIFGEPFLYLARPGKSNQDVALTYC 680
 QY 829 DLAKIHRDLLEVLNMPYFESDHFWSSLEITFNLDL-----NMLPGSGSELGSGFSR 883
 DB 829 DLAKIHRDLLEVLNMPYFESDHFWSSLEITFNLDL-----NMLPGSGSELGSGFSR 883
 QY 681 DLAKIHRDLLEVLNMPYFESDHFWSSLEITFNLDL-----NMLPGSGSELGSGFSR 738
 DB 681 DLAKIHRDLLEVLNMPYFESDHFWSSLEITFNLDL-----NMLPGSGSELGSGFSR 738
 QY 884 QRRKRLSPRRRTDKDTEQPGEVSAALGPGRAGAPSSRGHGPWGSESPSSGSS--SPSS 941
 DB 884 QRRKRLSPRRRTDKDTEQPGEVSAALGPGRAGAPSSRGHGPWGSESPSSGSS--SPSS 941
 QY 729 -----LSDNGSGSPHE--LGP-----QFPKGYSLILGPGSQ 767
 DB 729 -----LSDNGSGSPHE--LGP-----QFPKGYSLILGPGSQ 767
 QY 942 EDEGGRSSSPRLRYLFPSSPPRPGEPGCEPLMECEKSSSTCNPLSGAFSVSNIFESP 1001
 DB 942 EDEGGRSSSPRLRYLFPSSPPRPGEPGCEPLMECEKSSSTCNPLSGAFSVSNIFESP 1001
 QY 768 NSMGAG-----PCAPGHPDAAPL-----SISDASGLW 795
 DB 768 NSMGAG-----PCAPGHPDAAPL-----SISDASGLW 795
 QY 1002 GDSRROYOELPRCAPPTSLINPLSSPGRRPGDVE-----SRDLALOROLNRET 1054
 DB 1002 GDSRROYOELPRCAPPTSLINPLSSPGRRPGDVE-----SRDLALOROLNRET 1054
 QY 796 PE-----LLOEAP-----PRHSB-QSPQEDPOCWPLKLSRLQLOAOMNRLES 838
 DB 796 PE-----LLOEAP-----PRHSB-QSPQEDPOCWPLKLSRLQLOAOMNRLES 838

QY 1055 RUSADMAVYQLQLQROM-----TLVPPAVSAVTPPGD-----GPTS 1090
 DB 839 RVSSDLRILQQLQROMPGHASYLLEAPASNDLALVPIA-SETTSPGRPLPQGLPAPQ 897
 QY 1091 T-----SPLFVSPPLPTITLDSLOVQFMACELEP-----APELP 1128
 DB 896 TTSYDLDLDCSPKRRNSSFRMP--HLAAVTDKTYLAPSS--OQPEGLMPLASPLHP 951
 QY 1129 OE-----GPTRLSLPGQALGALTSQ-PLHRHSGDPG 1158
 DB 952 LEVQGLICGRCFS-SLPEHLISVPRKQLDQRRGSDPG 987

RESULT 13
 US-10-325-430-21
 ; Sequence 21, Application US/10325430
 ; Publication No. US20030153525A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Millennium Pharmaceuticals, Inc
 ; APPLICANT: Silos-Santiago, Immaculada
 ; APPLICANT: Rosenfeld, Julie Beth
 ; TITLE OF INVENTION: METHODS AND COMPOSITIONS IN TREATING
 ; TITLE OF INVENTION: PAIN AND PAINFUL DISORDERS USING 1465, 1587, 2146, 2207,
 ; FILE REFERENCE: MP101-294P1RM
 ; CURRENT APPLICATION NUMBER: US/10/325,430
 ; CURRENT FILING DATE: 2002-12-19
 ; PRIOR APPLICATION NUMBER: US 60/341,953
 ; PRIOR FILING DATE: 2001-12-19
 ; NUMBER OF SEQ ID NOS: 21
 ; SOFTWARE: FASTSEQ for Windows Version 4.0
 ; SEQ ID NO 21
 ; LENGTH: 958
 ; TYPE: PRF
 ; ORGANISM: Homo Sapiens
 US-10-325-430-21

Query Match 49.4%; Score 3001.5; DB 12; Length 958;
 Best Local Similarity 53.7%; Pred. No. 6.8e-210;
 Matches 659; Conservative 83; Mismatches 139; Indels 347; Gaps 30;

QY 1 MPVRGHVAPONTFLDTIIRKFEKGSRKFIANARVENCAYICNDGCELCGYSRAEYM 60
 DB 1 MPVRGHVAPONTFLDTIIRKFEKGSRKFIANARVENCAYICNDGCELCGYSRAEYM 60
 QY 61 GPRCTDFLHGRTQRRAAQIAOALLGAEEKVEIATFRKQSGCFCLVDVVPKKNEDG 120
 DB 61 GPRCTDFLHGRTQRRAAQIAOALLGAEEKVEIATFRKQSGCFCLVDVVPKKNEDG 120
 QY 121 AVIMEFLNEEVMKDMVGSPPADTNHRGPTSWLAPGRAKTFRLKLPALLA-LTARESS 179
 DB 121 AVIMEFLNEEVMKDMVGSPPADTNHRGPTSWLAPGRAKTFRLKLPALLA-LTARESS 179
 QY 180 VR-----SGAGAGAGAGAVVVDVLPAPAPSSSLALDEVYAMDNDHVAAGLGAPEERRA 233
 DB 180 VR-----SGAGAGAGAGAVVVDVLPAPAPSSSLALDEVYAMDNDHVAAGLGAPEERRA 233
 QY 146 QRLLSQSFSGSESGHGRPG----- 164
 DB 146 QRLLSQSFSGSESGHGRPG----- 164
 QY 234 LVGPSPRRAPGOLPSPRAHSLNPDASGSSCLARTSRRESCASVRRASADIEAMRA 293
 DB 234 LVGPSPRRAPGOLPSPRAHSLNPDASGSSCLARTSRRESCASVRRASADIEAMRA 293
 QY 165 --GPG-----PG----- 169
 DB 165 --GPG-----PG----- 169
 QY 294 GVLPPPPRHASTGAMHPLRSGLLNSTSDLVRYRTISKIPQITLNFVLDKGPFLASPT 353
 DB 294 GVLPPPPRHASTGAMHPLRSGLLNSTSDLVRYRTISKIPQITLNFVLDKGPFLASPT 353
 QY 170 -----TGR-----GKRTISOIQTPLNFVEFNLEKRRSST 201
 DB 170 -----TGR-----GKRTISOIQTPLNFVEFNLEKRRSST 201
 QY 354 SDRETIAP-KIKERTHNTEKTVQVLSIGADVLPYKLOAPRIHRWTILHYSPPKAVMDW 412
 DB 354 SDRETIAP-KIKERTHNTEKTVQVLSIGADVLPYKLOAPRIHRWTILHYSPPKAVMDW 412
 QY 202 TEIEIIPAKHYVERQNTVEKTVQVLSIGADVLPYKLOAPRIHRWTILHYSPPKAVMDW 261
 DB 202 TEIEIIPAKHYVERQNTVEKTVQVLSIGADVLPYKLOAPRIHRWTILHYSPPKAVMDW 261
 QY 413 LILLVITYAVFTPSAFLKETEGRPATCEGACQAPLAVVDLIVDIMEFVDILINR 472
 DB 413 LILLVITYAVFTPSAFLKETEGRPATCEGACQAPLAVVDLIVDIMEFVDILINR 472

```

QY 473 TTYVNAEEVYSHGRIAVHFKCMFLIDVAAIIPFDLJGSGSEE---LIGLTKARL 529
Db 321 TTYVNTDEVYSHRRRIAVHFKCMFLIDVAAIIPFDLJRTGSDETTTILGILTKARL 380
QY 530 LRLVAVARKLDYSEYGAAYVLLMCTCFALIAHMLACIWAIGNMDEQPHMDSRIGWHLNL 589
Db 381 LRLVAVARKLDYSEYGAAYVLLMCTCFALIAHMLACIWAIGNERERYLEHKGIMDSL 440
QY 590 GDQJGKPYNSG-LGQPSIKKKYVYALYFTTSSLTSGFGVNSPTNTEKTFSTCVMLIG 648
Db 441 GVOJGKRKNGSDPAGSGSVODKYVYALYFTTSSLTSGFGVNSPTNTEKTFSTCVMLIG 500
QY 649 SLMTASIFGNVAIIORLYSTAXYHOMLEVRREPIRHOIHPNPROLEXYFOHAWSYT 708
Db 501 SLMTASIFGNVAIIORLYSTAXYHOMLVKKEPIRHOIHPNPROLEXYFOHAWSYT 560
QY 709 NGIDMNAVLKGPCELOADICLHLNRSILOCHKEPFGATKLCRLALANKFKTTNAPGDT 768
Db 561 NGIDMNAVLKGPCELOADICLHLNRSILOCHKEPFGATKLCRLALANKFKTTNAPGDT 620
QY 769 LVHAGDILLALYFTRSGSIELTRDQVYVAILGKNDIGEPPLNTYARPKSGNDVRLATYC 828
Db 621 LVHAGDVLSTYFTRSGSIELTRDQVYVAILGKNDIGEPVSLHAQPKSSADVRLATYC 680
QY 829 DLKRIHRDDELLEVDMYPEFSDHWSLEITFNLRDT-----NMIPGSPGSTELEGCFSR 883
Db 681 DLKRIQRAADILEVDMYPAFAFESWPKLEVTFNLRDAAGLHSSRPARGQODHGF-- 738
QY 884 QRRKKLSFRRTDKDTEQPGEVSAIGPCAGAGFSSRKRGPGGPGWGESPPSSGSPSPESSED 943
Db 739 -----LSDNQSD-----AAPLSIDASGLW-----PELQOE 765
QY 944 EGPGRSSPLLVPEFSSRPGRPGGPGLMEDEKSSDTONPLSGATSGVSNIFSFYGD 1003
Db 766 -----MPPRHSPQSOED-----DC-----W-- 782
QY 1004 SRGROYELPRCPAPTPSLNIPILSPGRRRGDVESRLDALQROLNLETRLSDAMATV 1063
Db 783 -----PLK-----LGSLEQLQNMNLEBRVSDLSRI 811
QY 1064 LQLLQROH-----TLVPAYSAVTTPG-----GPTST----- 1091
Db 812 LQLLQKPMPOGHASYILEAPASNDLALVPIA-SETTSPGPRILPQGFILPPAOTPSYGLDD 870
QY 1132 PFRRLSLPGQALGALTSQ-PLHRHSGDP 1158
Db 925 PCFS-SLEPHLGSVPKQOLDPQRHSGSDPG 951

RESULT 14
US-10-192-440-2
: Sequence 2, Application US/10192440
: Publication No. US20030082718A1
: GENERAL INFORMATION:
: APPLICANT: Curtis, Roy A. J.
: TITLE OF INVENTION: 52908, A HUMAN POTASSIUM CHANNEL, AND
: FILE REFERENCE: MP12001-009PIRNM
: CURRENT APPLICATION NUMBER: US/10/192,440
: CURRENT FILING DATE: 2002-07-10
: PRIOR APPLICATION NUMBER: 60/341,953
: PRIOR FILING DATE: 2001-12-19
: PRIOR APPLICATION NUMBER: 60/304,243
: PRIOR FILING DATE: 2001-07-10
: NUMBER OF SEQ ID NOS: 15
: SOFTWARE: fastseq for Windows Version 4.0
: SEQ ID NO 2
: LENGTH: 958

```

Query Match	49.4%	Score 3001.5	DB 15	Length 958
Best Local Similarity	53.7%	Pred. No. 6,8e-210		
Matches	659	Conservative 83	Mismatches 139	Indels 347
			Gaps 30	
Query	1	MPVRGHA	PONTFLDITIRKFE	EGOSKRFTTANARVENC
Db	1	MPVRGHA	PONTYLDITIRKFE	EGOSKRFTTANAMENCAI
QY	61	QRPCTCDLHPR	QORRAAOLIA	ALGAEEKRIAYRRD
Db	61	QRPCTCDLHPR	QORRAAOLIA	ALGAEEKRIAYRRD
QY	121	AVIFILFE	VYVMEKDMVSP	ANDTNHNGPPTSWLP
Db	121	AVIFILFE	VYVMEKDMVSP	ANDTNHNGPPTSWLP
QY	180	VR-----	SGAGAGAG	AVYVDYDLTPA
Db	146	QRLTSQSF	LGSEGS	HGRPG-----
QY	224	LVGSGSP	PRSA	PGQPPRASHLN
Db	165	--GRP-----	PG-----	-----
QY	294	GVLPPR	PHASTGAMHPL	RSGILNSTSDSLV
Db	170	-----TGR-----	-----	-----
QY	354	SDRELI	AP-KIKERT	HNTEKVTQVLS
Db	202	TEIILIA	PHKVERFONTE	KVTOYLSIGADV
QY	413	LILLLV	ITAVFTPSA	AFILKEPNEP
Db	262	LILLLV	ITAVFTPSA	AFILKEPNEP
QY	473	TTYVANE	EVVSHPR	IAVHYEFKGF
Db	321	TTYVNT	NEVYSHPR	IAVHYEFKGF
QY	530	LRLV	VARAKLD	RYSEYGA
Db	381	LRLV	VARAKLD	RYSEYGA
QY	590	GDOIGK	PRVNSG	-LGSPIK
Db	441	GVOIGK	PRVNSG	-LGSPIK
QY	649	SLMTAS	IFGNVSA	IIORLYSGT
Db	501	SLMTAS	IFGNVSA	IIORLYSGT
QY	709	NGIDM	NAV	LKGP
Db	561	NGIDM	NAV	LKGP
QY	769	LVHAGD	LLALYF	SRGSI
Db	621	LVHAGD	LLALYF	SRGSI
QY	829	DLHK	IHRD	LLLEV
Db	681	DLHK	IHRD	LLLEV
QY	884	QRRK	LSFR	RRRDX
Db	739	-----	-----	-----
QY	944	EGPG	RSSP	LLVPESS

```

Db      766 -----MPPHSPQSPQEDP-----DC-----W--- 782
QY      1004 SGRQOYEIRCPAPPTSLNLTSSPGRRPGDVESRLDALROLNRLERLSADMATV 1063
Db      763 -----PLK-----LGRLEFOLOAOMNRLRESVSSDLSRI 811
QY      1064 LQLOLQOM-----TLVPPAYSAVTPPGP-----GPTST----- 1091
Db      812 LQLOLQPMPOGHASYILEAPASNDLALVPIA-SETTSPGRLPGGFLPAGTPSYGDLDD 870
QY      1092 -----SPLLPVPPLPTLTLDSLSOVQFMACELEPPG-----APELPG-----G 1131
Db      871 CSPKRRNSPPRMP--HLAVADTKTLPSE-----OEQPGMLPPLASPLHPLLEVQGLICG 924
QY      1132 PTRRLSLPGOLGALTSQ--PLHRGSDPG 1158
Db      925 PCFS-SLEPHLGSPVKQDLQFHRGSDPG 951

```

RESULT 15

```

US-10-192-440-9
; Sequence 9, Application US/10192440
; Publication No. US20030082718A1
; GENERAL INFORMATION:
; APPLICANT: Curtis, Rory A. J.
; TITLE OF INVENTION: 52908, A HUMAN POTASSIUM CHANNEL, AND
; TITLE OF INVENTION: USES THEREOF
; FILE REFERENCE: MP12001-009PIRNM
; CURRENT APPLICATION NUMBER: US/10/192,440
; CURRENT FILING DATE: 2002-07-10
; PRIOR APPLICATION NUMBER: 60/341,953
; PRIOR FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 60/304,243
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 950
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-10-192-440-9

```

```

Query Match      49.1%; Score 2982; DB 15; Length 950;
Best Local Similarity 52.9%; Pred. No. 1.8e-208;
Matches 647; Conservative 87; Mismatches 146; Indels 344; Gaps 25;

```

```

QY      1 MVRRGHVAPOMTFTLTIIRKFEQSRKFTIANAVENCAVYCCDGCCELCGYSRAEYM 60
Db      1 MVRRGHVAPOMTFTLTIIRKFEQSRKFTIANAVENCAVYCCDGCCELCGYSRAEYM 60
QY      61 QRPCTCDFLHGPRTOORAAQIAQALLGAEEKVEIAFYRKDGCFCFLADVVPYKNEDG 120
Db      61 QRPCTCDFLHGPRTOORAAQIAQALLGAEEKVEIAFYRKDGCFCFLADVVPYKNEDG 120
QY      121 AVIMEFLNEVYVMEKQVNGSPAHDTNHRGPTSWLAPGAKTFRLKLPALLALTARESSV 180
Db      121 AVIMEFLNEVYVMEKQVNGSPAHDTNHRGPTSWLAPGAKTFRLKLPALLALTARESSV 180
QY      121 AVIMEFLNEVYVMEKQVNGSPAHDTNHRGPTSWLAPGAKTFRLKLPALLALTARESSV 180
Db      121 AVIMEFLNEVYVMEKQVNGSPAHDTNHRGPTSWLAPGAKTFRLKLPALLALTARESSV 180
QY      181 RSGGAGGAGAPGAVVVDVLTLPAASSSESLADEVTAMDNHAGLGPAERALVGPSP 240
Db      181 RSGGAGGAGAPGAVVVDVLTLPAASSSESLADEVTAMDNHAGLGPAERALVGPSP 240
QY      241 PRSAPGQLPSPRAHSLNDPASGSSCLARTSRSCASVRRASADIEAMRAGVLPPPP 300
Db      241 PRSAPGQLPSPRAHSLNDPASGSSCLARTSRSCASVRRASADIEAMRAGVLPPPP 300
QY      174 ----- 173
QY      301 RHASTGAMHPLKSGLLNSTSDSLVRYRTSKIPQITLNFVDLKGDPFLASPTSDREIIA 360
Db      174 -----KYRTVSOIQPQFTLNFVEFNLEKHRSGSTTEIETIA 208
QY      361 P-KIKERTHNTEKTYOVLISGADVLPREKLOAPRIHRTIILHYSFKAVMWMLILLVI 419
Db      361 P-KIKERTHNTEKTYOVLISGADVLPREKLOAPRIHRTIILHYSFKAVMWMLILLVI 419

```

```

Db      209 PAKVVERQNTYEKTYOVLISGADVLPREKLOAPRIHRTIILHYSFKAVMWMLILLVI 268
QY      420 YTAVFPTYSAAFLKETECEGPATECGYACOPAVVDLIVDIMFIVDLINERTTYVAN 479
Db      269 YTAVFPTYSAAFLKETECEGPATECGYACOPAVVDLIVDIMFIVDLINERTTYVAN 479
QY      480 EEVVSHPGRIAVHYKGNFLIDMVAALFPDILLFGSSGE--LIGLKTARLLRVYA 536
Db      328 DEVVSHPRRIAVHYKGNFLIDMVAALFPDILLFGSSGE--LIGLKTARLLRVYA 387
QY      537 RKLDRYSEGAIVFLMLCTFALLNHWLACIYATGAGNNEOPMDSRIGMLHLDQIGKP 596
Db      368 RKLDRYSEGAIVFLMLCTFALLNHWLACIYATGAGNNEOPMDSRIGMLHLDQIGKP 447
QY      597 YNSG-IGGPSIKDKYVVALYFTFSSLSVSGVNSPMTNSEKIPISICVMLIGSLMAYSI 655
Db      448 YNSGDPASGPSVQDKYVVALYFTFSSLSVSGVNSPMTNSEKIPISICVMLIGSLMAYSI 507
QY      656 FGNVSAIIQRLYSGTARHTQMLRYREFIRHQIPIPLRQLEEFQHAMSTYNGIDMNA 715
Db      508 FGNVSAIIQRLYSGTARHTQMLRYREFIRHQIPIPLRQLEEFQHAMSTYNGIDMNA 567
QY      716 VLKGFPECLQADICHLNRSLLOHCKPPRGATKGLRALAMKFKTHAPPGDTLVHADL 775
Db      568 VLKGFPECLQADICHLNRSLLOHCKPPRGATKGLRALAMKFKTHAPPGDTLVHADL 627
QY      776 LTAIFYISRSIEILRGDVVAALCKNDIFGEPLMLYARPGKSGNDVVALTYCDLHKIHR 835
Db      628 LSTLYFISRSIEILRGDVVAALCKNDIFGEPLMLYARPGKSGNDVVALTYCDLHKIHR 687
QY      836 DDLEVLDMYEPESCHFWSSLETFNLDNTN---MTP-GSPGSTELEGGSROKRRKLS 890
Db      688 ADLEVLDMYPAFAOTFENKLEVFENLMDADGLOSTPROAPGHDPGQFF----- 738
QY      891 FRRRTDKTDEQGEVSALGPRGAGPSSRGPRGPGPWGESPPSSPSSSEDEBPGRSS 950
Db      739 -----LNDSSGAPs----- 749
QY      951 SPLRLVFPSSPRPGEPPGEPIMDECKESDPTCNPLSGAFSGVNSIFPMGDSNGRQYO 1010
Db      750 -----LSTSDTSALW----- 759
QY      1011 ELPRCPATPPLL-NIPSSPGRRRRGVES-----RLDAIQOLNRLERLSADMAT 1062
Db      760 -----PELLQMPSPSP--NPRODLDCWHRELGFRLQLOAOMNRLRESVSSDLSR 808
QY      1063 VLQLOLQOMTLVPPAY-----SAVTPPGGP-----TSTPPL----- 1095
Db      809 ILQLOLQHPQG--RPSYILGASASDLSASFPEISVTRSSSESTLVGHVSAQTLISGDLDD 866
QY      1096 -----PVSPLPTLTLDSLSOVQFMACELEPPGAPPELPQEGPTRLSLPGQLA-- 1144
Db      867 HIGTRFNPSRPRPHVAMAMDKTLVPS-----EQEGPGGLSPLASPLRPLEVPGIGSRF 922
QY      1145 -----LTSQP-----LHRGSDPG 1158
Db      923 SLPEHLSVSPKOLEFORHSDPG 946

```

Search completed: September 23, 2003, 17:48:51
 Job time : 463 secs